GenCore version 5.1.6 Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2005, 15:38:47; Search time 8472.11 Seconds

(without alignments)

12117.322 Million cell updates/sec

Title: US-10-624-932-1 COPY 46 2742

Perfect score: 2697

Sequence: 1 atggccgtccggcct.....tgtcggaggctgagtgctga 2697

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 segs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_htc:*

4: gb est3:*

5: gb est4:*

6: gb est5:*

7: gb_est6:*

8: gb qss1:*

9: gb gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		8			•	
Result		Query				
No.	Score	Match	Length	DB	ID	Description
1	960.2	35.6	2802	9	AY406491	AY406491 Homo sapi
2	950.4	35.2	2791	9	AY406493	AY406493 Mus muscu
3	923.4	34.2	3790	3	AK031655	AK031655 Mus muscu
4	874	32.4	1852	3	CR598115	CR598115 full-leng
5	871.4	32.3	3866	3	AK018177	AK018177 Mus muscu
6	814	30.2	2802	9	AY406492	AY406492 Pan trogl
7	810.4	30.0	2532	9	AY411747	AY411747 Homo sapi
8	780.4	28.9	2532	9	AY411749	AY411749 Mus muscu

	9	768.6	28.5	1034	4	BI758231	BT758231	603029876
	10	736.6	27.3	1532	3	BC033727		Homo sapi
	11	735.2	27.3	788	1	AI951556		wv36f04.x
С	12	721.6	26.8	796	5	BX348193		BX348193
C	13	721.0	26.5	818	4	BI818609		603033362
	14	678.2	25.1	2532	9	AY411748		Pan trogl
	15	672.6	24.9	934	2	BF311804		601897316
	16	665	24.7	859	2	BF311896		601897733
	17	647.4	24.0	853	5	BX364574		BX364574
	18	621	23.0	900	5	BX345406		BX345406
	19	618.8	22.9	756	5	BU612387		UI-M-EWO-
	20	610.2	22.6	2775	9	AY401471		Mus muscu
	21	603.2	22.4	874	5	BQ689148		AGENCOURT
	22	600.6	22.4	889	5	BQ691915		AGENCOURT
	23	600.2	22.3	977	5	BX345407		BX345407
	23 24	599	22.3	2775	9	AY401469		Homo sapi
		595.2	22.2	601				AL516580
	25					AL516580		601147261
	26	589.4	21.9	604	2	BE314370		
	27	586.6	21.8	1072	5	BX422753		BX422753
	28	566.6	21.0	678	6	CA749784		UI-M-FDO-
	29	561	20.8	1175	2	BF530640		602071931
	30	555.6	20.6	788	6	CA317532		UI-M-FWO-
	31	555	20.6	572	7	CR554569		DKFZp459I
	32	491.2	18.2	499	7	CR747398		CR747398
	33	486.2	18.0	548	5	BX452510		BX452510
	34	471.2	17.5	540	4	BI467799		389583 MA
	35	470.8	17.5	2507	9	AY401470		Pan trogl
	36	468.6	17.4	824	4	BI737024		603360874
	37	438.2	16.2	675	6	CA315487		UI-M-FW0-
	38	433	16.1	471	5	BX282095		BX282095
	39	428.4	15.9	1147	5	BU840446		AGENCOURT
	40	424	15.7	843	4	BG913440	BG913440	602811321
	41	417.8	15.5	856	7	CN164143	CN164143	994266 MA
	42	411.8	15.3	749	7	CF735417		UI-M-HBO-
	43	410.6	15.2	1111	4	BG298307		602397080
	44	409.8	15.2	751	7	CF735550		UI-M-HBO-
	45	408.8	15.2	460	2	BF443156	BF443156	260542 MA

ALIGNMENTS

RESULT 1

```
AY406491
                                     2802 bp
LOCUS
                                                        linear
                                                                 GSS 15-DEC-2003
            AY406491
                                                DNA
            Homo sapiens UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION
            genomic survey sequence.
            AY406491
ACCESSION
VERSION
            AY406491.1 GI:39762465
KEYWORDS
            GSS.
SOURCE
            Homo sapiens (human)
  ORGANISM
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 2802)
REFERENCE
  AUTHORS ·
            Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
            Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
```

```
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
          Adams, M.D. and Cargill, M.
 TITLE
          Inferring nonneutral evolution from human-chimp-mouse orthologous
          gene trios
          Science 302 (5652), 1960-1963 (2003)
 JOURNAL
  PUBMED
          14671302
REFERENCE
          2 (bases 1 to 2802)
 AUTHORS
          Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
          Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
          Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
          Adams, M.D. and Cargill, M.
 TITLE
          Direct Submission
 JOURNAL
          Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
          Rockville, MD 20850, USA
          This sequence was made by sequencing genomic exons and ordering
COMMENT
          them based on alignment.
                  Location/Qualifiers
FEATURES
                  1. .2802
    source
                  /organism="Homo sapiens"
                  /mol type="genomic DNA"
                  /db xref="taxon:9606"
                  <1. .>2802
    gene
                  /gene="UNC5C"
                  /locus tag="HCM2575"
ORIGIN
 Query Match
                       35.6%;
                              Score 960.2; DB 9;
                                                 Length 2802;
 Best Local Similarity
                       61.3%; Pred. No. 1.9e-194;
 Matches 1682; Conservative
                             0; Mismatches 995;
                                                         69;
                                                                    6;
                                                 Indels
                                                             Gaps
         12 GCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGG 71
Qy
                        1
                            Db
         66 GCTCGTGCTACCTGCCCTGGCCCTGCTCAGCGCCAGCGGCACTGGCTCCGCCGCCCAAGA 125
         72 TGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCC 131
Qy
                                1
                              1
                                                            11 11
Db
        126 TGATGACTTTTTCATGAACTCCCAGAAACTTTTCCTTCTGATCCACCTGAGCCTCTGCC 185
        132 CCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGT 191
Qy
             186 ACATTTCCTTATTGAGCCTGAAGAAGCTTATATTGTGAAGAATAAGCCTGTGAACCTGTA 245
Db
        192 GTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCA 251
Qv
                        246 CTGTAAAGCAAGCCCTGCCACCCAGATCTATTTCAAGTGTAATAGTGAATGGGTTCATCA 305
Db
        252 GGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGT 311
Qу
               \Pi
                                                            306 GAAGGACCACATAGTAGATGAAAGAGTAGATGAAACTTCCGGTCTCATTGTCCGGGAAGT 365
Db
        312 CCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTG 371
Qу
             366 GAGCATTGAGATTTCGCGCCAGCAAGTGGAAGAACTCTTTGGACCTGAAGATTACTGGTG 425
Db
        372 CCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCAT 431
Qу
            Db
        426 CCAGTGTGTGGCCTGGAGCTCCGCGGGTACCACAAAGAGCCGGAAGGCGTATGTGCGCAT 485
```

QĀ	432	AGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCT	485
Db	486		545
Qу	486	GGAGCAGGGCATCGTGCCGTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGA	545
Db	546	GGAACAGGAAGTCTTACTCCAGTGTCGACCACCTGAAGGGATCCCAGTGGCTGAGGTGGA	605
Qу	546	GTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCG	605
Db	606	ATGGTTGAAAAATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTTATATTACTAT	665
Qу	606	GGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGT	665
Db	666	TGATCACAACCTCATCATAAAGCAGGCCCGACTCTCTGATACTGCAAATTACACCTGTGT	725
Qу	666	GGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTCATCGTCTACGTGAA	725
Db	726	TGCCAAAAACATTGTTGCCAAGAGGAAAAGTACAACTGCCACTGTCATAGTCTATGTCAA	785
Qy	726	CGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTG	785
Db	786	CGGTGGCTGGTCCACCTGGACGGAGTGGTCTGTGTAACAGCCGCTGTGGACGAGGGTA	845
Qy	786	GCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGA	845
Db	846	TCAGAAACGTACAAGGACTTGTACCAACCCGGCACCACTCAATGGGGGTGCCTTCTGTGA	905
Qy	846	GGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAG	905
Db	906	AGGGCAGAGTGTGCAGAAAATAGCCTGTACTACGTTATGCCCAGTGGATGGCAGGTGGAC	965
Qу	906	CCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGACTGCACCCACTGGCGGAGCCGTGAGTG	965
Db	966	GCCATGGAGCAAGTGGTCTACTTGTGGAACTGAGTGCACCCACTGGCGCAGGAGGGAG	1025
Qу	966	CTCTGACCCAGCACCCGCAACGGAGGGGAGGGGGGGGGG	1025
Db	1026	CACGGCGCCAGCCCCAAGAATGGAGGCAAGGACTGCGACGGCCTCGTCTTGCAATCCAA	1085
Qу	1026	CAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTA	1085
Db	1086	GAACTGCACTGATGGGCTTTGCATGCAGACTGCTCCTGATTCAGATGATGTTGCTCTCTA	1145
Qу	1086	TGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTGTCCTCATCCT	1142
Db	1146	TGTTGGGATTGTGATAGCAGTGATCGTTTGCCTGGCGATCTCTGTAGTTGTGGCCTTGTT	1205
Qу	1143	CGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCAC	1202
Db	1206	TGTGTATCGGAAGAATCATCGTGACTTTGAGTCAGATATTATTGACTCTTCGGCACTCAA	1265
Qу	1203	CTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCAC	1262
Db -	1266	TGGGGGCTTTCAGCCTGTGAACATCAAGGCAGCAAGACAAGA	1316

Qу	1263	CATCCAGCCGGACCTCAGCACCACCACCACCAGGGCAGTCTCTGTCCCCGGCA	1322
Db	1317	TGTACCCCCAGACCTCACGTCAGCTGCAGCCATGTACAGAGGACCTGTCTATGCCCTGCA	1376
Qу	1323	GGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCC	1373
Db	1377	TGACGTCTCAGACAAAATCCCAATGACCAACTCTCCAATTCTGGATCCACTGCCCAACCT	1436
Qу	1374	CCTGGGTGGCGCCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTT	1433
Db	1437	GAAAATCAAAGTGTACAACACCTCAGGTGCTGTCACCCCCCAAGATGACCTCTCTGAGTT	1496
Qу	1434	CGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACAT	1493
Db	1497	TACGTCCAAGCTGTCCCCTCAGATGACCCAGTCGTTGTTGGAGAATGAAGCCCTCAGCCT	1556
Qу	1494	GACCTATGGGACCTTCAA	1511
Db		GAAGAACCAGAGTCTAGCAAGGCAGACTGATCCATCCTGTACCGCATTTGGCAGCTTCAA	
Qу	1512	CTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGA	1571
Db	1617	CTCGCTGGGAGGTCACCTTATTGTTCCCAATTCAGGAGTCAGCTTGCTGATTCCCGCTGG	1676
Qу	1572	TGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGT	1631
Db	1677	GGCCATTCCCCAAGGGAGAGTCTACGAAATGTATGTGACTGTACACAGGAAAGAAA	1736
Qy	1632	GAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCC	1691
Db	1737	GAGGCCACCCATGGATGACTCTCAGACACTTTTGACCCCTGTGGTGAGCTGTGGGCCCCC	1796
Qу	1692	TGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCC	1751
Db	1797	AGGAGCTCTGCTCACCCGCCCAGTCGTCCTCACTATGCATCACTGCGCAGACCCCAATAC	1856
Qу	1752	TGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCT	1811
Db	1857	CGAGGACTGGAAAATACTGCTCAAGAACCAGGCAGCACAGGGACAGTGGGAGGATGTGGT	1916
Qy	1812	GCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTG	1871
Db	1917	GGTGGTCGGGGAGAAAACTTCACCACCCCTGCTACATTCAGCTGGATGCAGAGGCCTG	1976
Qу	1872	CTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGC	1931
Db	1977	CCACATCCTCACAGAGAACCTCAGCACCTACGCCCTGGTAGGACATTCCACCACAAAGC	2036
QУ	1932	TGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC	1991
Db	2037	GGCTGCGAAGCGCCTCAAGCTGGCCATCTTTGGGCCCCTGTGCTGCTCCTCGCTGGAGTA	2096
Qу	1992	CAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCT	2051
Db	2097	CAGCATCCGAGTCTACTGTCTGGATGACACCCAGGATGCCCTGAAGGAAATTTTACATCT	2156
Qу	2052	GGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAG	2111

Db	2157	
Qу	2112	TTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCT 2171
Db	2217	
Qу	2172	CCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTT 2231
Db	2277	GCTGGCTAAATATCAGGAAATTCCATTTTACCATGTTTGGAGTGGATCTCAAAGAAACCT 2336
Qу	2232	GCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCT 2291
Db	2337	GCACTGCACCTTCACTCTGGAAAGATTTAGCCTGAACACAGTGGAGCTGGTTTGCAAACT 2396
Qy	2292	GTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAA 2351
Db	2397	CTGTGTGCGGCAGGTGGAAGGAGAAGGGCAGATCTTCCAGCTCAACTGCACCGTGTCAGA 2456
Qу	2352	GGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGT 2411
Db	2457	GGAACCTACTGGCATCGATTTGCCGCTGCTGGATCCTGCGAACACCATCACCACGGTCAC 2516
Qу	2412	GGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGA 2471
Db	2517	GGGGCCCAGTGCTTTCAGCATCCCTCTCCCTATCCGGCAGAAGCTCTGTAGCAGCCTGGA 2576
Qу	2472	CCCACCTGTAGGCGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAG 2531
Db	2577	TGCCCCCAGACGAGAGGCCATGACTGGAGGATGCTGGCCCATAAGCTGAACCTGGACAG 2636
QУ	2532	CCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGA 2591
Db		GTACTTGAATTACTTTGCCACCAAATCCAGCCCAACTGGCGTAATCCTGGATCTTTGGGA 2696
Qу		GGCGCGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTG
Db		AGCACAGAACTTCCCAGATGGAAACCTGAGCATGCTGGCAGCTGTCTTGGAAGAAATGGG 2756
Qy -		CCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
Db	2757	AAGACATGAAACGGTGGTCCTTAGCAGCAGAAGGGCAGTATTAA 2802
RESULT 2 AY406493 LOCUS		Y406493 2791 bp DNA linear GSS 15-DEC-2003
DEFINITI		us musculus UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence, enomic survey sequence.
ACCESSIO VERSION	N Ā	Y406493 Y406493.1 GI:39762467
KEYWORDS SOURCE		SS. us musculus (house mouse)
ORGANI		us musculus ukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENC		ammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. (bases 1 to 2791)

```
AUTHORS
          Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
          Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
          Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
          Adams, M.D. and Cargill, M.
 TITLE
          Inferring nonneutral evolution from human-chimp-mouse orthologous
          gene trios
 JOURNAL
          Science 302 (5652), 1960-1963 (2003)
  PUBMED
          14671302
REFERENCE
            (bases 1 to 2791)
 AUTHORS
          Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
          Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
          Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
          Adams, M.D. and Cargill, M.
          Direct Submission
 TITLE
          Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 JOURNAL
          Rockville, MD 20850, USA
          This sequence was made by sequencing genomic exons and ordering
COMMENT
          them based on alignment.
                  Location/Oualifiers
FEATURES
                  1. .2791
    source
                  /organism="Mus musculus"
                  /mol type="genomic DNA"
                  /db xref="taxon:10090"
                  <1. .>2791
    gene
                  /gene="UNC5C"
                  /locus tag="HCM2575"
ORIGIN
 Query Match
                       35.2%;
                              Score 950.4; DB 9;
                                                 Length 2791;
                       62.3%; Pred. No. 2.3e-192;
 Best Local Similarity
 Matches 1653; Conservative
                             0; Mismatches 921;
                                                 Indels
                                                         80;
                                                             Gaps
                                                                    7;
        104 TGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACA 163
Qу
                         11 11 1
                                 Db
        158 TTCCATCTGACCCACCTGAGCCATTGCCACACTTCCTCATTGAGCCCGAGGAAGCTTACA 217
        164 TCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCT 223
Qу
            1 11 1111111111
                                - 11
                                      Db
        224 TCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACG 283
Qγ
            278 TCAAGTGCAACAGCGAGTGGGTTCATCAGAAGGACCACGTAGTAGACGAGAGAGTAGATG 337
Db
        284 GGAGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGA 343
Qу
              338 AAACCTCTGGTCTAATTGTGAGAGAGTGAGCATTGAGATTTCACGCCAGCAGGTGGAGG 397
Db
        344 AGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCA 403
Qy
               398 AACTGTTTGGGCCTGAAGATTACTGGTGCCAGTGTGTGGCCTGGAGCTCAGCAGGCACTA 457
Db
         404 CCAAGAGTCAGAAGGCCTACATCCGCATAGCCAG--
                                            ----ATTGCGCAAGAACTTCGAGC 457
Qу
            4414 4111 11111
         458 CGAAGAGTCGGAAGGCATACGTGCGCATTGCGTTTCCAGATCTGCGGAAGACATTCGAGC 517
Db
        458 AGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCCGTCCCTGCCGTCCAC 517
Qv
```

Db	518	AGGAACCCTTGGGAAAGGAAGTGTCCTTGGAGCAGGAAGTCTTACTCCAGTGTCGGCCAC	577
Qy	518	CGGAGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGT	577
Db	578	CTGAAGGGATCCGTAGAATGGCTAAAGAATGAAGACATAATTGATCCTG	626
Qy	578	CCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCC	637
Db	627	CTGAAGATCGGAACTTTTATATTACTATCGATCACAACCTGATCATCAAGCAAG	686 ⁻
Qу	638	TTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCG	697
Db	687	TCTCAGATACAGCAAATTATACCTGTGTTGCCAAAAATATTGTTGCCAAGAGAAAAAAGCA	746
Qу	698	CCTCCGCTGTCATCGTCTACGTGAACGGTGGGTCGACGTGGACCGAGTGGTCCG	757
Db ,	747	CCACAGCCACTGTCATCGTGTATGTTAATGGTGGCTGGTCCACCTGGACAGAGTGGTCTG	806
Qу	758	TCTGCAGCGCCAGCTGTGGGCGGCGGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGG	817
Db	807	TGTGTAACAGCCGCTGTGGGCGAGGATATCAGAAACGCACAAGAACCTGCACCAACCCAG	866
Qy .	818	CGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCA	877
Db	867	CCCCACTCAATGGTGGGGCCTTCTGTGAGGGGCAGAGTGTGCAGAAAATAGCATGCACTA	926
Qу	878	CCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGG	937
Db	927	CGTTATGTCCAGTGGATGGTAGGTGGACTTCATGGAGCAAATGGTCAACCTGTGGGACTG	986
Qу	938	ACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGG	997
Db	987	AATGCACCCACTGGCGCAGGAGGGAGTGTACAGCACCAGCCCCCAAGAACGGGGGTAAGG	1046
Qу	998	AGTGCCAGGGCACTGACCTGGACACCGCAACTGTACCAGTGACCTCTGTGTACACAGTG	1057
Db	1047	ACTGTGATGGCCTGGTCCTCCAATCCAAGAACTGCACTGATGGGCTGTGCATGCA	1106
Qy	1058	CTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCC	1114
Db	1107	CTCCTGACTCAGATGATGTGGCTCTCTACGTGGGGATTGTGATCGCTGTAACAGTCTGTC	1166
Qу	1115	TGGTCCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACT	1174
Db	1167	TGGCGATCACTGTTGTGTGGCCCTGTTTGTGTATCGGAAGAACCACCGTGACTTTGAGT	1226
Qу	1175	CAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCA	1234
Db	1227	CTGACATCATTGACTCCTCAGCACTCAATGGCGGCTTTCAGCCTGTGAACATCAAG	1282
Qу	1235	GCAAAGCAGACACCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCA	1294
Db	1283	GCTGCCAGACAAGATCTCCTGGCTGTCCCCCCTGACCTCACCTCAGCTGCAGCCA	1337
Qу	1295	CCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCT	1349

Db	1338	TGTACAGGGGACCTGTCTATGCTCTGCATGATGTCTCAGACAAAATCCCAATGACCAACT 1397	
Qу	1350	CACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTC 1409	
Db	1398	CTCCAATTCTGGACCCACTACCCAACTTGAAAATCAAAGTGTACAACAGCTCAGGTGCTG 1457	
Qу	1410	TCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCC 1453	
Db	1458	TCACTCCTCAGGATGACCTTGCCGAGTTCTCATCCAAACTGTCACCCCAGATGACCCAGT 1517	
Qу	1454	AGAACTACTTCCGCTCCCTGCCCCGAGGCA 1483	
Db	1518	CCTTGCTAGAGAATGAGGCCCTTAACCTGAAGAACCAGAGCCTCGCAAGACAGAC	
Qу	1484	CCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATA 1543	
Db	1578	CATCCTGCACAGCATTTGGTACCTTCAACTCTCTTGGGGGTCACCTCATCATTCCTAATT 1637	
Qу	1544	CAGGTATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCT 1603	
Db	1638	CAGGAGTAAGCTTGCTGATTCCCGCTGGGGCCATTCCTCAGGGGAGAGTCTATGAAATGT 1697	
Qу	1604	ACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTG	
Db	1698	ATGTGACTGTACACAGGAAAGAAATATGAGGCCCCCCATGGAAGACTCTCAGACCCTAC 1757	
Qу	1664	TGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTCACCCGGCCAGTCATCCTGG 1723	
Db	1758	TTACCCCTGTGGTGAGCTGTGGGCCTCCTGGAGCTCTGCTGACCCGCCCTGTCATCCTCA 1817	
Qу	1724	CTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGT 1783	
Db	1818	CTCTGCATCACTGTGCAGACCCCAGCACCGAGGACTGGAAGATCCAGCTCAAAAACCAGG 1877	
Qу		CGTGCGAGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCT 1843	
Db	1878	CAGTGCAGGGACAATGGGAGGATGTTGTGGTGGTTGGGGAGGAGAACTTCACAACCCCCT 1937	
Qу	1844	ACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTG 1903	
Db	1938	GTTACATTCAGCTGGATGCAGAGGCTTGCCATATCCTCACAGAGAACCTCAGTACCTATG 1997	
Qу	1904	CCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTG 1963	
Db	1998	CCCTGGTTGGGCAGTCCACCACCAAAGCAGCTGCCAAGCGTCTTAAACTGGCCATCTTTG 2057	
QУ	1964	CGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	
Db	2058	GGCCCCTCTGCTGCTCTTCCCTGGAGTACAGCATTAGAGTCTACTGCCTGGATGACACAC 2117	
QУ	2024	ACGATGCACTCAAGGAGGTGCTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGG 2083	
Db		AGGATGCCCTGAAGGAAGTTCTACAACTGGAGAGGCAAATGGGAGGACAGCTCCTAGAAG 2177	
Qу		AGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	
Db	2178	AACCCAAGGCTCTTCATTTTAAAGGCAGCATCCACAACCTGCGCCTGTCTATTCATGACA 2237	

Qy		TGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATC 2203
Db	2238	TCGCCCATTCCCTCTGGAAGAGCAAATTGCTGGCTAAGTATCAGGAAATTCCATTTTACC 2297
Qy	2204	ACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCC 2263
Db	2298	ACATCTGGAGTGGCTCTCAAAGAAACCTCCACTGCACCTTCACTCTGGAAAGACTCAGCC 2357
Qу	2264	CCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGA 2323
Db	2358	TAAACACAGTGGAACTGGTTTGCAAACTCTGTGTGCGGCAGGTTGAAGGAGAAGGGCAGA 2417
Qy	2324	GCTTCAGCATCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGG 2383
Db	2418	
Qу	2384	AGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCA 2443
Db	2478	ACCCTGCTAGTACCATCACCACTGTCACCGGACCAAGTGCTTTCAGCATTCCTCTCCCTA 2537
Qу	2444	TTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGA 2503
Db	2538	
Qy	2504	CTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCC 2563
Db	2598	
Qу	2564	CCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCC 2623
Db	2658	
Qу	2624	AGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGACGCTGGCCTCTTCACAGTGTCGG 2683
Db ·	2718	
Qy	2684	AGGCTGAGTGCTGA 2697
Db	2778	
RESULT 3 AK031655 LOCUS DEFINITI	A ON M e	K031655 3790 bp mRNA linear HTC 03-APR-2004 us musculus 13 days embryo male testis cDNA, RIKEN full-length nriched library, clone:6030473H24 product:unc5 homolog (C.
ACCESSIO VERSION KEYWORDS SOURCE ORGANI	N A H M SM M E	<pre>legans) 3, full insert sequence. K031655 K031655.1 GI:26327502 TC; CAP trapper. us musculus (house mouse) us musculus ukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; ammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.</pre>
REFERENC		

AUTHORS

Carninci, P. and Hayashizaki, Y.

```
High-efficiency full-length cDNA cloning
  TITLE
            Meth. Enzymol. 303, 19-44 (1999)
  JOURNAL
            99279253
 MEDLINE
            10349636
   PUBMED
REFERENCE
            2
            Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 AUTHORS
            Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
            Normalization and subtraction of cap-trapper-selected cDNAs to
  TITLE
            prepare full-length cDNA libraries for rapid discovery of new genes
            Genome Res. 10 (10), 1617-1630 (2000)
  JOURNAL
            20499374
 MEDLINE
   PUBMED
            11042159
REFERENCE
            3
            Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 AUTHORS
            Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
            Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
            Yamamoto, R., Matsumoto, H., Sakaquchi, S., Ikegami, T., Kashiwaqi, K.,
            Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
            Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
            Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE
            RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL
            Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE
            20530913
   PUBMED
            11076861
REFERENCE
 AUTHORS
            The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
 TITLE
            Functional annotation of a full-length mouse cDNA collection
  JOURNAL
            Nature 409, 685-690 (2001)
REFERENCE
            5
 AUTHORS
            The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
 TITLE
            Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL
            Nature 420, 563-573 (2002)
REFERENCE
               (bases 1 to 3790)
 AUTHORS
            Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
            Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
            Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
            Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
            Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
            Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
            Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
            Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
            Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
            Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
            Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
            Muramatsu, M. and Hayashizaki, Y.
 TITLE
            Direct Submission
  JOURNAL
            Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
            URL: http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
```

```
COMMENT
            cDNA library was prepared and sequenced in Mouse Genome
           Encyclopedia Project of Genome Exploration Research Group in Riken
           Genomic Sciences Center and Genome Science Laboratory in RIKEN.
           Division of Experimental Animal Research in Riken contributed to
           prepare mouse tissues.
           Please visit our web site for further details.
           URL:http://genome.gsc.riken.jp/
           URL:http://fantom.gsc.riken.jp/.
FEATURES
                    Location/Qualifiers
                     1. .3790
    source
                     /organism="Mus musculus"
                     /mol type="mRNA"
                     /strain="C57BL/6J"
                     /db xref="FANTOM DB:6030473H24"
                     /db xref="taxon:10090"
                     /clone="6030473H24"
                     /sex="male"
                     /tissue type="testis"
                     /clone lib="RIKEN full-length enriched mouse cDNA library"
                     /dev stage="13 days embryo"
    CDS
                     118. .2970
                     /note="unnamed protein product; putative
                    unc5 homolog (C. elegans) 3 (MGD|MGI:1095412,
                    GB|NM 009472, evidence: BLASTN, 99%, match=464)"
                     /codon start=1
                     /protein id="BAC27495.1"
                     /db xref="GI:26327503"
                     translation="MRKGLRATAARCGLGIGYLLQMLVLPALALLSASGTGSAAQDDE/
                     FFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQ
                     KDHVVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYV
                    RIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEDIIDPAEDRNFYI
                     TIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRC
                     GRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIACTTLCPVDGRWTSWSKWSTCGTECTH
                    WRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQGFIYPISTEHRPQNEYGFSSAPD
                     SDDVALYVGIVIAVTVCLAITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAA
                     RQDLLAVPPDLTSAAAMYRGPVYALHDVSDKIPMTNSPILDPLPNLKIKVYNSSGAVT
                     PODDLAEFSSKLSPOMTOSLLENEALNLKNOSLAROTDPSCTAFGTFNSLGGHLIIPN
                     SGVSLLI PAGAI PQGRVYEMYVTVHRKENMRPPMEDSQTLLTPVVSCGPPGALLTRPV
                     ILTLHHCADPSTEDWKIQLKNQAVQGQWEDVVVVGEENFTTPCYIQLDAEACHILTEN
                     LSTYALVGQSTTKAAAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQM
                     GGQLLEEPKALRFKGSIHNLRLSIHDIAHSLWKSKLLAKYQEIPFYHIWSGSQRNLHC
                     TFTLERLSLNTVELVCKLCVRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPASTITTVT
                     GPSAFSIPLPIRQKLCSSLDAPQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDL
                    WEAQNFPDGNLSMLAAVLEEMGRHETVVYLAAEGQY"
ORIGIN
 Query Match
                          34.2%;
                                 Score 923.4; DB 3;
                                                      Length 3790;
 Best Local Similarity
                         61.3%; Pred. No. 1.4e-186;
 Matches 1659; Conservative
                                 0; Mismatches 926;
                                                      Indels 120;
                                                                             6:
         104 TGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACA 163
Qy
                            11111
                                     \mathbf{1}
Db
         275 TTCCATCTGACCCACCTGAGCCATTGCCACACTTCCTCATTGAGCCCGAGGAAGCTTACA 334
```

164 TCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCT 223

Qу

Db

Qу	224	TCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACG	283
Db	395	TCAAGTGCAACAGCGAGTGGGTTCATCAGAAGGACCACGTAGTAGACGAGAGAGTAGATG	454
Qу	284	GGAGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGA	343
Db	455	AAACCTCTGGTCTAATTGTGAGAGAAGTGAGCATTGAGATTTCACGCCAGCAGGTGGAGG	514 ,
Qу	344	AGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCA	403
Db	515	AACTGTTTGGGCCTGAAGATTACTGGTGCCAGTGTGTGGCCTGGAGCTCAGCAGGCACTA	574
Qу	404	CCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGC	463
Db	575	CGAAGAGTCGGAAGGCATACGTGCGCATTGCGTATCTGCGGAAGACATTCGAGCAGGAAC	634
Qу	464	CGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGG	523
Db	635	CCTTGGGAAAGGAAGTCTTGGAGCAGGAAGTCTTACTCCAGTGTCGGCCACCTGAAG	694
Qу	524	GCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGG	583
Db	695	GGATCCCAGTGGCTGAGGTAGAATGGCTAAAGAATGAAGACATAATTGATCCTGCTGAAG	754
Qу	584	ACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTG	643
Db	755	ATCGGAACTTTTATATTACTATCGATCACAACCTGATCATCAAGCAAG	814
Qу	644	ACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCG	703
Db	815	ATACAGCAAATTATACCTGTGTTGCCAAAAATATTGTTGCCAAGAGAAAAAAGCACCACAG	874
Qу	704	CTGCTGTCATCGTCTACGTGAACGGTGGGTCGACGTGGACCGAGTGGTCCGTCTGCA	763
Db	875	CCACTGTCATCGTGTATGTTAATGGTGGCTGGTCCACCTGGACAGAGTGGTCTGTGTA	934
Qy	764	GCGCCAGCTGTGGGCGGCGCGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTC	823
Db ·	935	ACAGCCGCTGTGGGCGAGGATATCAGAAACGCACAAGAACCTGCACCAACCCAGCCCCAC	994
Qy	824	TCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGT	883
Db	995	TCAATGGTGGGGCCTTCTGTGAGGGGCAGAGTGTGCAGAAAATAGCATGCACTACGTTAT	1054
Qу	884	GCCCAGTAGACGGCAGCTGGAGCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCA	943
Db	1055		1114
Qу	944	CCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCC	1003
Db	1115		1174
Qу	1004	AGGGCACTGACCTGGACACCGCAACTGTACCAGTGACCTCTGTGTACACAG	1055
Db	1175		1234

QУ	1056TGCTTCTGGCC 1066
Db	1235 ACCCCATTTCAACTGAGCACAGACCCCAGAATGAATATGGATTTTCTTCTGCTCCTGACT 1294
Qу	1067 CTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGC 1123
Db	1295 CAGATGATGTGGCTCTCTACGTGGGGATTGTGATCGCTGTAACAGTCTGTCT
Qу	1124 TGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGG 1183
Db	1355 CTGTTGTGGTGGCCCTGTTTGTGTATCGGAAGAACCACCGTGACTTTGAGTCTGACATCA 1414
Qу	1184 CTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAG 1243
Db	1415 TTGACTCCTCAGCACTCAATGGCGGCTTTCAGCCTGTGAACATCAAGGCTG 1465
Qу	1244 ACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCACCAGG 1303
Db	1466 CCAGACAAGATCTCCTGGCTGTCCCCCCTGACCTCACCTCAGCTGCAGCCATGTACAGGG 1525
QУ	1304 GCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGG 1358
Db	1526 GACCTGTCTATGCTCTGCATGATGTCTCAGACAAAATCCCAATGACCAACTCTCCAATTC 1585
Qу	1359 GCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACTGCACCACAGCTCTCCCACCTC 1418
Db	1586 TGGACCCACTACCCAACTTGAAAATCAAAGTGTACAACAGCTCAGGTGCTGTCACTCCTC 1645
Qy	1419TGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCC 1453
Db	1646 AGGATGACCTTGCCGAGTTCTCATCCAAACTGTCACCCCAGATGACCCAGTCCTTGCTAG 1705
Qy	1454AGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACA 1492
Db	1706 AGAATGAGGCCCTTAACCTGAAGAACCAGAGCCTCGCAAGACAGAC
QУ	1493 TGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCA 1552
 Db	1766 CAGCATTTGGTACCTTCAACTCTCTTGGGGGTCACCTCATCATTCCTAATTCAGGAGTAA 1825
QУ	1553 GCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGC 1612
Db	1826 GCTTGCTGATTCCCGCTGGGGCCATTCCTCAGGGGAGAGTCTATGAAATGTATGT
QУ	1613 TGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTG
Db ~	1886 TACACAGGAAAGAAATATGAGGCCCCCCATGGAAGACTCTCAGACCCTACTTACCCCTG 1945
Qу	1673 TCGTTAGCTGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACC 1732
Db	1946 TGGTGAGCTGTGGGCCTCCTGGAGCTCTGCTGACCCGCCCTGTCATCCTCACTCTGCATC 2005
QУ	1733 ACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGG 1792
Db	2006 ACTGTGCAGACCCCAGCACCGAGGACTGGAAGATCCAGCTCAAAAACCAGGCAGTGCAGG 2065
Qy	1793 GCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCC 1852

Db	2066		2125
Qу	1853	AGCTGGAGGCCAGTGCCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGG	1912
Db	2126		2185
Qу	1913	GAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGG	1972
Db	2186		2245
Qу	1973	CCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2032
Db	2246		2305
QУ	2033	TCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGG	2092
Db	2306		2365
Qу	2093	TCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	2152
Db	2366	CTCTTCGTTTTAAAGGCAGCATCCACAACCTGCGCCTGTCTATTCATGACATCGCCCATT	2425
Qy	2153	CCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGA	2212
Db	2426	CCCTCTGGAAGAGCAAATTGCTGGCTAAGTATCAGGAAATTCCATTTTACCACATCTGGA	2485
QУ	2213	ATGGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTA	2272
Db	2486	GTGGCTCTCAAAGAAACCTCCACTGCACCTTCACTCTGGAAAGACTCAGCCTAAACACAG	2545
. Qу	2273	GTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCA	2332
Db	2546	TGGAACTGGTTTGCAAACTCTGTGTGCGGCAGGTTGAAGGAGAAGGGCAGATCTTCCAGC	2605
Qу	2333	TCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAG	2392
Db	2606	TCAACTGTACTGTCAGAGGAACCTACTGGCATCGACTTACCTCTCCTGGACCCTGCTA	2665
Qy	2393	CGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGA	2452
Db	2666	GTACCATCACCACTGTCACCGGACCAAGTGCTTTCAGCATTCCTCTCCCTATCCGGCAGA	2725
Qy		AGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCC	2512
Db		AGCTATGCAGCAGCCTGGATGCCCCTCAAACAAGAGGCCATGACTGGAGGATGCTGGCCC	2785
QУ	2513	AGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCA	2572
Db	2786	ATAAACTCAACCTGGACAGGTACTTGAATTACTTTGCCACCAAATCGAGCCCAACTGGCG	2845
QУ	2573	TGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGC	2632
Db	2846	TAATCCTGGATCTTTGGGAAGCACAGAACTTCCCAGATGGAAACCTGAGCATGCTGGCAG	2905
Qу	2633	CAGCAGTGGCTGGACTGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGT	2692

```
Db
        2906 CCGTCCTGGAAGAATGGGAAGACATGAGACAGTGGTGTACTTGGCAGCAGAAGGACAGT 2965
        2693 GCTGA 2697
Qу
               111
Db
        2966 ATTGA 2970
RESULT 4
CR598115
                                                             HTC 21-JUL-2004
                                            mRNA
                                                    linear
LOCUS
           CR598115
                                  1852 bp
           full-length cDNA clone CS0DA006YG16 of Neuroblastoma of Homo
DEFINITION
           sapiens (human).
ACCESSION
           CR598115
VERSION
           CR598115.1 GI:50478922
KEYWORDS
           HTC; CNSLT cDNA.
SOURCE
           Homo sapiens (human)
  ORGANISM
           Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
           1 (bases 1 to 1852)
 AUTHORS
           Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
  TITLE
           Full-length cDNA libraries and normalization
           Unpublished
  JOURNAL
  REMARK
           Contact : Feng Liang Email : fliang@lifetech.com URL :
           http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
           Faraday Avenue
              (bases 1 to 1852)
REFERENCE
           2
  AUTHORS
           Genoscope.
  TITLE
           Direct Submission
  JOURNAL
           Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage:
           BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
           - Web : www.genoscope.cns.fr)
COMMENT
           1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
           end enriched, double-strand cDNA was digested with Not I and cloned
           into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
           was normalized. Library was constructed by Life Technologies, a
           division of Invitrogen.
FEATURES
                    Location/Oualifiers
                    1. .1852
    source
                    /organism="Homo sapiens"
                    /mol type="mRNA"
                    /db xref="taxon:9606"
                    /clone="CS0DA006YG16"
                    /tissue type="Neuroblastoma"
                    /plasmid="pCMVSPORT_6"
ORIGIN
  Query Match
                         32.4%; Score 874; DB 3; Length 1852;
  Best Local Similarity
                         100.0%; Pred. No. 4.8e-176;
 Matches 874; Conservative
                               0; Mismatches
                                                0; Indels
                                                              0;
                                                                  Gaps
                                                                          0;
        1824 GGAGGCGCCCTCCCACCTCTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCAC 1883
Qу
             Db
           1 GGAGGCGCCCTCCCACCTCTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCAC 60
        1884 CGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCG 1943
Qy
```

Db	61	$\tt CGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCG$	120
Qу	1944	CCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC	2003
Db	121	CCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC	180
Qу	2004	CTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCT	2063
Db	181	CTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCT	240
Qу	2064	GGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCT	2123
Db	241	GGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCT	300
Qу	2124	GCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTA	2183
Db	301	GCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTA	360
Qу	2184	CCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTT	2243
Db	361	CCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACCTT	420
Qy	2244	CACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCA	2303
Db	421	CACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCA	480
Qу	2304	GGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTT	2363
Db	481	GGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACATCACCAAGGACACAAGGTT	540
Qу	2364	TGCTGAGCTGCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGC	2423
Db	541	TGCTGAGCTGCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGC	600
Qу	2424	CTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCC	2483
Db	601	CTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCC	660
Qу	2484	GCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTT	2543
Db	661	GCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTT	720
Qу	2544	CTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGCACTT	2603
Db	721	CTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGCACTT	780
Qу	2604	CCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCA	2663
Db	781	CCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCA	840
Qу	2664	TGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697	
Db	841	TGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 874	

```
AK018177
LOCUS -
                                     3866 bp
                                                 mRNA
                                                          linear
                                                                   HTC 03-APR-2004
            Mus musculus adult male medulla oblongata cDNA, RIKEN full-length
DEFINITION
            enriched library, clone:6330415E02 product:TRANSMEMBRANE RECEPTOR
            UNC5H2 homolog [Rattus norvegicus], full insert sequence.
ACCESSION
            AK018177
            AK018177.1 GI:12857775
VERSION
KEYWORDS
            HTC; CAP trapper.
            Mus musculus (house mouse)
SOURCE
            Mus musculus
  ORGANISM
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
            Carninci, P. and Hayashizaki, Y.
 AUTHORS
            High-efficiency full-length cDNA cloning
  TITLE
            Meth. Enzymol. 303, 19-44 (1999)
  JOURNAL
            99279253
 MEDLINE
            10349636
   PUBMED
REFERENCE
            2
  AUTHORS
            Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
            Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
  TITLE
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
            Genome Res. 10 (10), 1617-1630 (2000)
  JOURNAL
            20499374
 MEDLINE
            11042159
   PUBMED
REFERENCE
  AUTHORS
            Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
            Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
            Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
            Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
            Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
            Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
            Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
            RIKEN integrated sequence analysis (RISA) system--384-format
  TITLE
            sequencing pipeline with 384 multicapillary sequencer
            Genome Res. 10 (11), 1757-1771 (2000)
  JOURNAL
 MEDLINE
            20530913
            11076861
   PUBMED
REFERENCE
 AUTHORS
            The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
            Functional annotation of a full-length mouse cDNA collection
  TITLE
  JOURNAL
            Nature 409, 685-690 (2001)
REFERENCE
            5
 AUTHORS
            The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE
            Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL
            Nature 420, 563-573 (2002)
                (bases 1 to 3866)
REFERENCE
  AUTHORS
            Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
            Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
            Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
            Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
            Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
            Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
```

Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,

Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216) Please visit our web site (http://genome.gsc.riken.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B. Location/Qualifiers 1. .3866 source /organism="Mus musculus" /mol type="mRNA" /strain="C57BL/6J" /db xref="FANTOM DB:6330415E02" /db xref="taxon:10090" /clone="6330415E02" /sex="male" /tissue type="medulla oblongata" /clone lib="RIKEN full-length enriched mouse cDNA library" /dev stage="adult" 417. .3254 /note="unnamed protein product; TRANSMEMBRANE RECEPTOR UNC5H2 homolog [Rattus norvegicus] (SPTR|008722, evidence: FASTY, 96.5%ID, 100%length, match=2835) putative" /codon start=1 /protein id="BAB31108.1" /db xref="GI:12857776" translation="MRARSGVRSALLLALLLCWDPTPSLAGVDSAGQVLPDSYPSAPA/ EQLPYFLLEPQDAYIVKNKPVELHCRAFPATQIYFKCNGEWVSQNDHVTQESLDEATG LRVREVQIEVSRQQVEELFGLEDYWCQCVAWSSSGTTKSRRAYIRIAYLRKNFDQEPL AKEVPLDHEVLLOCRPPEGVPVAEVEWLKNEDVIDPAQDTNFLLTIDHNLIIRQARLS DTANYTCVAKNIVAKRRSTAATVIVYVNGGWSSWAEWSPCSNRCGRGWQKRTRTCTNP

> APLNGGAFCEGQAFQKTACTTVCPVDGAWTEWSKWSACSTECAHWRSRECMAPPPQNG GRDCSGTLLDSKNCTDGLCVLNQRTLNDPKSHPLETSGDVALYAGLVVAVFVVVAVLM

TITLE

COMMENT

FEATURES

CDS

JOURNAL

AEGVIVYRRNCRDFDTDITDSSAALTGGFHPVNFKTARPNNPQLLHPSAPPDLTASAG IYRGPVYALQDSADKIPMTNSPLLDPLPSLKIKVYNSSTIGSGSGLADGADLLGVLPP GTYPGDFSRDTHFLHLRSASLGSQHLLGLPRDPSSSVSGTFGCLGGRLSLPGTGVSLL VPNGAIPQGKFYDLYLHINKAESTLPLSEGSQTVLSPSVTCGPTGLLLCRPVVLTVPH CAEVIAGDWIFQLKTQAHQGHWEEVVTLDEETLNTPCYCQLEAKSCHILLDQLGSYVF MGESYSRSAVKRLQLAIFAPALCTSLEYSLRVYCLEDTPVALKEVLELERTLGGYLVE EPKPLLFKDSYHNLRLSLHDIPHAHWRSKLLAKYQEIPFYHVWNGSQRALHCTFTLER HSLASTEFTCKVCVRQVEGEGQIFQLHTTLAETPAGSLDALCSAPGÑAITTQLGPYAF KIPLSIRQKICSSLDAPDSRGNDWRLLAQKLSMDRYLNYFATKASPTGVILDLWEARQ QDDGDLNSLASALEEMGKSEMLVAMATDGDC"

ORIGIN

	cal S	32.3%; Score 871.4; DB 3; Length 3866; Similarity 60.6%; Pred. No. 1.9e-175;	
Matches	1645	5; Conservative 0; Mismatches 941; Indels 129; Gaps	8;
Qy	112	GCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAG	171
Db	540	GCCCTGCGGAGCAGCTGCCGTACTTCCTATTGGAGCCACAGGACGCCTACATCGTAAAG	599
Qу	172	AACAAGCCAGTGCTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGC	231
Db	600	AACAAGCCAGTGGAACTGCACTGCAGAGCCTTCCCTGCCACGCAGATCTACTTCAAGTGT	659
Qy	232	AACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGT	291
Db	660	AATGGCGAGTGGGTCAGCCAGAATGACCACGTCACACAGGAGAGCCTGGATGAGGCCACA	719
Qу	292	GGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTC	351
Db	720	GGCTTGCGGGTGCGAGAGTGCAGATCGAGGTGTCACGGCAGCAAGTGGAGGAACTCTTC	779
Qy	352	GGGCTGGAGGAATAÇTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGT	411
Db	780	GGGCTCGAGGACTACTGGTGCCAGTGCGTGGCCTGGAGCTCTTCGGGAACTACCAAGAGT	839
Qy	412	CAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCC	471
Db	840		899
Qу	472	AAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCGTCCACCGGAGGGCATCCCT	531
Db	900	AAGGAGGTACCCTTGGATCATGAGGTCCTTCTGCAGTGCCGCCCACCGGAGGGAG	959
Qy	532	CCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAAT	591
Db	960	GTGGCTGAGGTGGAATGGCTCAAGAATGAAGATGTCATTGACCCCGCTCAGGACACTAAC	1019
Qy	592	GTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCC	651
Db	1020	TTCCTGCTCACCATTGACCACAACCTCATCATCCGCCAGGCGCGCCTCTCAGACACGGCC	1079
Qy	652	AACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTC	711
Db	1080	AACTACACCTGTGTGGCCAAGAATATCGTGGCCAAGCGCCGGAGCACCGCGGCCACAGTC	1139
Qу	712	${\tt ATCGTCTACGTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGC}$	771

Db 1140		1199
Qy 772	TGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGG	831
Db 1200	TGTGGCCGAGGCTGCAGAAGCGTACTCGGACCTGCACCAATCCAGCCCCACTCAATGGA	1259
Qy 832	GGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTA	891
Db 1260	GGCGCCTTCTGTGAGGGACAGGCCTTCCAGAAGACAGCTTGCACCACCGTGTGCCCAGTG	1319
Qy 892	GACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGG	951
Db 1320	GATGGAGCGTGGACCGAGTGGAGCAAGTGGTCTGCCTGCAGCACAGAGTGTGCGCACTGG	1379
Qy 952	CGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACT	1011
Db 1380	CGCAGCCGCGAGTGCATGGCACCGCCACCCCAGAACGGAGGCCGTGACTGCAGCGGGACG	1439
Qy 1012	GACCTGGACACCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCC	1066
Db 1440	CTACTTGACTCCAAGAACTGCACTGATGGGCTGTGCGTGC	1499
Qy 1067	CTGAGGACGTGGCCCTCTATGTGGGCCTCATC	1098
Db 1500	GACCCTAAAAGCCACCCCTGGAGACATCGGGAGATGTGGCACTGTACGCAGGCCTTGTG	1559
Qy 1099	GCCGTGGCCGTCTGCCTGCTGCTGCTGCTGTCCTCATCCTCGTTTATTGCCGG	1155
Db 1560	GTGGCCGTCTTTGTGGTGGTAGCGGTTCTCATGGCCGAGGGAGTGATCGTATACCGGAGA	1619
Qy 1156	AAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC	1212
Db 1620	AACTGCCGGGACTTCGACACGGACATCACCGACTCCTCTGCGGCCCTCACTGGTGGCTTC	1679
Qy 1213	CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATC	1266
Db 1680	CACCCTGTCAACTTCAAGACTGCAAGGCCCAACAACCCGCAGCTCCTGCACCCGTCCGCC	1739
	CAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGA-	1325
	CCTCCAGACCTAACGGCCAGTGCTGGCATCTACCGCGGGCCTGTGTATGCCCTGCAGGAC	1799
Qy 1326	TGGGCCCAGCCCCAAG	1341
Db 1800	TCCGCCGACAAGATCCCCATGACTAATTCGCCCCTGCTGGATCCCCTGCCCAGCCTCAAG	1859
Qy 1342	TTCCAGCTCACCAATGGGCACCTGCTCAGC	1371
Db 1860	ATCAAGGTCTATAACTCCAGCACCATCGGTTCTGGGTCTGGCCTGATGGAGCCGAC	1919
Qy 1372	CCCCTGGGTGGCGGCCGCCACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAG	1431
Db 1920	CTGCTGGGTGTCCTCCCGCCGGGCACGTACCCAGGCGATTTCTCCCGGGACACCCATTTC	1979
Qy 1432	TTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACC	1485

Db	1980	CTGCACCTGCGCAGTGCCAGCCTTGGTTCCCAGCACCTCCTGGGCCTACCTCGGGACCCC	2039
Qу	1486	AGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACA	1545
Db	2040		2099
Qy	1546	GGTATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTAC	1605
Db	2100		2159
Qу	1606	CTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTG	1665
Db	2160	CTACATATCAACAAGGCCGAAAGCACCCTCCCACTTTCAGAAGGTTCCCAGACAGTATTG	2219
Qу	1666	AGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCT	1725
Db	2220	AGCCCTCGGTGACCTGTGGGCCCACAGGCCTACTCCTGTGCCGCCCTGTCGTCCTCACC	2279
Qу	1726	ATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCG	1785
Db	2280	GTGCCCCACTGTGCTGAAGTCATCGCTGGAGACTGGATCTTTCAGCTCAAGACCCAGGCC	2339
Qy	1786	TGCGAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTAC	1845
Db	2340	CATCAGGGCCACTGGGAGGAGGTGGTGACCTTGGATGAGGAGACCCTCAACACCCCTGC	2399
Qу	1846	TACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCC	1905
Db	2400	TACTGCCAGCTGGAGGCTAAGTCCTGCCACATCCTGCTGGACCAGCTGGGTTCCTACGTA	2459
Qу	1906	CTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCG	1965
Db	2460	TTCATGGGCGAGTCCTACTCTCGCTCTGCAGTCAAGCGGCTCCAGCTGGCCATCTTCGCC	2519
Qу	1966	CCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2025
Db	2520	CCAGCCCTCTGCACCTCCCTGGAGTATAGCCTCAGGGTCTACTGTCTGGAGGACACACCT	2579
Qy		GATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAG	2085
Db	2580	${\tt GTAGCACTGAAGGAGGTCCTGGAGGAGGAGCTCTGGGTGGCTACTTGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG$	2639
Qy	2086	CCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	2145
Db	2640	CCCAAGCCTTTGCTCTTTAAGGACAGTTACCACAACCTACGCCTCTCCCTCC	2699
Qy	2146	CCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCAC	2205
Db		CCCCATGCCCACTGGAGGAGCAAACTACTGGCCAAGTACCAGGAGATTCCCTTCTACCAC	
Qу	2206	ATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCC	2265
Db	2760	GTCTGGAATGGCAGCCAGAGAGCCCTGCACTTCACCCTGGAGAGGCATAGCCTG	2819
Qу	2266	AGCACTAGTGACCTGCCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGC	2325
Dh	2820	$CCCTCC\Delta CGC\Delta CTTC\Delta CCTCT\Delta \Delta CGTCTCCCCTCCC$	2879

```
2326 TTCAGCATCAACTTCAACATCACC---AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTG 2382
Qy
                           1
                               1 11
                                       1 1
                                             1 1 1
                                                             11
             111
                   1 11
                                                        11
        2880 TTCCAGCTGCACACACGTTGGCCGAGACGCCTGCTGGCTCCCTGGATGCTCTCTGCTCT 2939
Db
        2383 GAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTC 2442
Qy
                        2940 GCCCGGGCAATGCCATCACCAGCTGGGACCCTATGCCTTCAAGATACCCCTGTCC 2999
Db
        2443 ATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGG 2502
Qу
            3000 ATCCGCCAAAAGATCTGCAGCAGCCTGGACGCCCCGACTCCCGGGGCAACGACTGGAGG 3059
Db
        2503 ACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGC 2562
Qv
                               Db
        3060 CTGTTGGCGCAGAAGCTGTCCATGGACCGGTACCTAAACTACTTCGCCACCAAAGCTAGT 3119
        2563 CCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGC 2622
Qv
                      3120 CCCACAGGTGTCATCTTAGACCTCTGGGAAGCTCGGCAACAGGATGACGGGGACCTCAAC 3179
Db
        2623 CAGCTGGCTGCAGCAGTGGCTGGACTGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCG 2682
Qу
                11111
                        1.1
                           - 1
                                     -11
                                                            11
        3180 AGCCTGGCCAGTGCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTAGCCATGGCCACA 3239
Db
        2683 GAGGCTGAGTGCTGA 2697
Qу
             11 | 11 | 11 | 11 | 11
        3240 GATGGCGATTGCTGA 3254
Db
RESULT 6
AY406492
                                  2802 bp
                                                            GSS 15-DEC-2003
LOCUS
           AY406492
                                            DNA
                                                    linear
           Pan troglodytes UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION
           genomic survey sequence.
ACCESSION
           AY406492
           AY406492.1 GI:39762466
VERSION
KEYWORDS
SOURCE
           Pan troglodytes (chimpanzee)
  ORGANISM
           Pan troglodytes
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
              (bases 1 to 2802)
REFERENCE
           Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
  AUTHORS
           Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
           Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
           Adams, M.D. and Cargill, M.
           Inferring nonneutral evolution from human-chimp-mouse orthologous
 TITLE
           gene trios
           Science 302 (5652), 1960-1963 (2003)
  JOURNAL
  PUBMED
           14671302
REFERENCE
             (bases 1 to 2802)
 AUTHORS
           Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
           Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
           Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
           Adams, M.D. and Cargill, M.
  TITLE
           Direct Submission
```

```
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
   JOURNAL
                    Rockville, MD 20850, USA
                    This sequence was made by sequencing genomic exons and ordering
COMMENT
                    them based on alignment.
                                    Location/Qualifiers
FEATURES
                                    1. .2802
        source
                                    /organism="Pan troglodytes"
                                    /mol type="genomic DNA"
                                    /db xref="taxon:9598"
                                    <1. .>2802
         gene
                                    /gene="UNC5C"
                                    /locus tag="HCM2575"
ORIGIN
                                                          Score 814; DB 9; Length 2802;
                                             30.2%;
   Query Match
                                             53.8%; Pred. No. 3.4e-163;
   Best Local Similarity
                                                                                                               69; Gaps
                                                         0; Mismatches 1200; Indels
   Matches 1477; Conservative
                   12 GCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGG 71
Qу
                                                      1
                   66 GCTCGTGCTACCTGCCCTGGCCCTGCTCAGCGCCAGCGGCACTGGCTCCGCCGCCCAAGA 125
Db
                   72 TGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCC 131
Qy
                                                                                                        11 11 1 11 11
                                                          - 1
                                             126 TGATGACTTTTTCATGAACTCCCAGAAACTTTTCCTTCGGATCCACCTGAGCCTCTGCC 185
 Db
                  132 CCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGT 191
 Qу
                           186 ACATTTCCTTATTGAACCTGAGGAAGCTTATATTGTGAAGAATAAGCCTGTGAACCTGTA 245
 Db
                  192 GTGCAAGGCC-----GTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGT 245
 Qy
                           11 11 11
                  Db
                  246 GCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCAT 305
 Qу
                  Db
                  306 GGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATA 365
 Qy
                  366 ИПИНИТИТЕТ В ТОТОТИТЕТ В Т
 Db
                  366 CTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCAACAAGAGTCAGAAGGCCTACAT 425
 Qy
                  Db
                   426 CCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCT 485
 Qy
                                                     11 1111 11 1111111 11 11 1 1 1111 11111
                   Db
                   486 GGAGCAGGGCATCGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGA 545
 Qy
                                            546 GGAACAGGAAGTCTTACTCCAGTGTCGGCCACCTGAAGGGATCCCAGTGGCTGAGNNNNN 605
 Db
                   546 GTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCG 605
 Qу
                                                                                         41 1 411 1 11 11 11
                                           13 41 111 1 1 11 11
                   606 NNNNNGAAAAATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTTATATTACTAT 665
```

Db

	QУ	606	GGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGT	665
	Db	666	TGATCACAACCTCATCATAAAGCAGGCCCGACTCTCTGATACTGCAAATTATACCTGTGT	725
	Qу	666	GGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAA	725
	Db	726	TGCCAAAAACATTGTTGCCAAGAGGAAAAGTACAACTGCCACTGTCATAGTCTATGTCAA	785
	Qy	726	CGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTG	785
	Db	786	CGGTGGCTGGTCCACCTGGACGGAGTGGTCTGTGTGTAACAGCCGCTGTGGACGAGGGTA	845
	Qy	786	GCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGA	845
	Db	846	TCAGAAACGTACAAGGACTTGTACCAACCCGGCACCACTCAATGGGGGTGCCTTCTGTGA	905
	Qу	846	GGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAG	905
	Db	906	AGGGCAGAGTGTGCAGAAAATAGCCTGTACTACGTTATGCCCAGTGGATGGCAGGTGGAC	965
	Qy	906	CCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGACTGCACCCACTGGCGGAGCCGTGAGTG	965
	Db	966	GCCATGGAGCAAGTGGTCTACTTGTGGAACTGAGTGCACCCACTGGCGCAGGAGGGAG	1025
	Qу	966	CTCTGACCCAGCACCCGCAACGGAGGGGAGGGGGGGGGG	1025
	Db	1026	CACGGCGCCAGCCCCCAAGAATGGAGGCAAGGACTGCGATGGCCTCGTCTTGCAATCCAA	1085
	Qу	1026	CAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTA	1085
	Db	1086	GAACTGCACTGATGGGCTTTGCATGCAGACTGCTCCTGATTCAGATGATGTTGCTCTCTA	1145
	Qу	1086	TGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTGCTCATCCT	1142
	Db	1146	TGTTGGGATTGTGATAGCAGTGATCGTTTGCCTGGCGATCTCTGTAGTTGTGGCCTTGTT	1205
	Qy ·	1143	CGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCAC	1202
	Db	1206	TGTGTATCGGAAGAATCATCGTGACTTTGAGTCAGATATTATTGACTCTTCGGCACTCAA	1265
	Qy	1203	CTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCAC	1262
	Db	1266	TGGGGGCTTTCAGCCTGTGAACATCAAGGCAGCAAGACAAGA	1316
•	Qу	1263	CATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCCCGGCA	1322
	Db	1317	TGTACCCCCAGACCTCACGTCAGCTGCAGCCATGTACAGAGGACCTGTCTATGCCCTGCA	1376
	Qу	1323	GGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCC	1373
	Db	1377	TGACGTCTCAGACAAAATCCCAATGACCAACTCTCCAATTCTGGATCCACTGCCCAACCT	1436
	Qу	1374	CCTGGGTGGCGCCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTT	1433
	Db	1437	GAAAATCAAAGTGTACAACACCTCAGGTGCTGTCACCCCCCAAGATGACCTCTCTGAGTT	1496

Qy	1434	CGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACAT	1493
Db	1497	TACGTCCAAGCTGTCCCCTCAGATGACCCAGTCATTGTTGGAGAATGAAGCCCTCAGCCT	1556
Qу	1494	GACCTATGGGACCTTCAA	1511
Db	1557	GAAGAACCAGAGTCTAGCAAGGCAGACTGATCCATCCTGTACCGCATTTGGCAGCTTCNN	1616
Qy	1512	CTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGA	1571
Db	1617	NTNGCTGGGNNNNCACCTTATTGTTCCCAATTCAGGAGTCAGCTTGCTGATTCCCGCTGG	1676
Qу	1572	TGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGT	1631
Db	1677	GGCCATTCCCCAAGGGAGAGTCTACGAAATGTATGTGACTGTACACAGGAAAGAAA	1736
Qy	1632	GAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCC	1691
Db	1737	GAGGCCACCCATGGATGACTCTCAGACACTTTTGACCCCTGTGGTGAGCTGTGGGCCCCC	1796
Qy	1692	TGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCC	1751
Db	1797		1856
Qу	1752	TGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCT	1811
Db	1857		1916
Qy	1812	GCACCTGGGCGAGGAGGCCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTG	1871
Db	1917	GGTGGTCGGGGAGAAACTTCACCACCCCTGCTACATTCAGCTGGATGCAGAGGCCTG	1976
Qy	1872	CTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGC	1931
Db	1977	CCACATCCTCACAGAGAACCTCAGCACCTACGCCCTGGTAGGACATTCCACCACAAAGC	2036
Qу	1932	TGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC	1991
Db	2037		2096
Qу	1992	CAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCT	2051
Db	2097	CAGCATCCGAGTCTACTGTCTGGATGACACCCAGGATGCCCTGAAGGAAATTTTACATCT	2156
Qy	2052	GGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAG	2111
Db	2157	TGAGAGACAGATGGGAGGACAGCTCCTAGAAGAACCTAAGGCTCTTCATTTTAAAGGCAG	2216
Qy	2112	TTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCT	2171
Db	2217	CACCCACAACCTGCGCCTGTCAATTCACGATATCGCCCATTCCCTCTGGAAGAGCAAATT	2276
Qy	2172	CCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTT	2231
Db	2277		2336
Qy	2232	GCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCT	2291

```
11 1111 11111 11
Db
        2337 GCACTGCACCTTCACTCTGGAAAGATTTAGCCTGAACACAGTGGAGCTGGTTTGCAAACT 2396
Qy
        2292 GTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAA 2351
             11 111 111111111 11 11 11 11111 1111
                                                 2397 CTGTGTGCGGCAGGTGGAAGGAGAAGGGCAGATCTTCCAGCTCAACTGCACCGTGTCAGA 2456
Db
        2352 GGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGT 2411
Qу
            111 11
                          11 11 1 1111 11
                                                      1111111
        2457 GGAACCTACTGGCATCGATTTGCCGCTGCTGGATCCTGCGAACACCATCACCACGGTCAC 2516
Db
        2412 GGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGA 2471
Qу
            1 11111111
Db
        2517 GGGGCCCAGTGCTTTCAGCATCCCTCTCCCTATCCGGCAGAAGCTCTGTAGCAGCCTGGA 2576
        2472 CCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAG 2531
Qу
              \mathbf{I}
                    \perp
                                Db
        2532 CCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGGGA 2591
Qy
                2637 GTACTTGAATTACTTTGCCACCAAATCCAGCCCAACTGGCGTAATCCTGGATCTTTGGGA 2696
Db
        2592 GGCGCGCACTTCCCCAACGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGG 2651
Qy
             2697 AGCACAGAACTTCCCAGATGGAAACCTGAGCATGCTGGCAGCTGTCTTGGAAGAAATGGG 2756
Db
Qу
        2652 CCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
                1 11 1 1
                          1 11 1
                                              1 1
        2757 AAGACATGAAACGGTGGTGTCCTTAGCAGCAGAAGGGCAGTATTAA 2802
Db
RESULT 7
AY411747
LOCUS
          AY411747
                                2532 bp
                                         DNA
                                                 linear
                                                         GSS 12-DEC-2003
DEFINITION
          Homo sapiens HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,
          genomic survey sequence.
          AY411747
ACCESSION
VERSION
          AY411747.1 GI:39767715
KEYWORDS
          GSS.
SOURCE
          Homo sapiens (human)
 ORGANISM
          Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
             (bases 1 to 2532)
          Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 AUTHORS
          Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
          Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
          Adams, M.D. and Cargill, M.
 TITLE
          Inferring nonneutral evolution from human-chimp-mouse orthologous
          gene trios
          Science 302 (5652), 1960-1963 (2003)
 JOURNAL
  PUBMED
          14671302
REFERENCE
          2 (bases 1 to 2532)
 AUTHORS
          Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
          Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
          Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
```

```
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 JOURNAL
         Rockville, MD 20850, USA
         This sequence was made by sequencing genomic exons and ordering
COMMENT
         them based on alignment.
FEATURES
                Location/Qualifiers
                1. .2532
    source
                /organism="Homo sapiens"
                 /mol type="genomic DNA"
                /db xref="taxon:9606"
                <1. .>2532
    gene
                /locus tag="HCM4327"
ORIGIN
                    30.0%;
                           Score 810.4; DB 9;
 Query Match
                                           Length 2532;
 Best Local Similarity
                    61.1%; Pred. No. 2e-162;
 Matches 1540; Conservative
                          0; Mismatches 851;
                                                  129;
                                           Indels
                                                       Gaps
                                                             9;
        307 GAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATAC 366
Qy
           13 GAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTCTTTGGGCTGGAGGATTAC 72
Db
        367 TGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCAAGAGTCAGAAGGCCTACATC 426
Qу
           111111 11
Db
        73 TGGTGCCAGTGCGTGGCCTGGAGCTCCGCGGGCACCAAGAGTCGCCGAGCCTACGTC 132
        427 CGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTG 486
Qy
                     133 CGCATCGCCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTGGGCAAGGAGGTGCCCCTG 192
Db
        487 GAGCAGGGCATCGTGCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAG 546
Qу
                     1 111
                  193 GACCATGAGGTTCTCCTGCAGTGCCGCCGGCGGAGGGGGGTGCCTGTGGCCGAGGTGGAA 252
Db
        547 TGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGG 606
Qy
                253 TGGCTCAAGAATGAGGATGTCATCGACCCCACCCAGGACACCAACTTCCTGCTCACCATC 312
Db
        607 GAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTG 666
Qу
           313 GACCACAACCTCATCATCCGCCAGGCCCGCCTGTCGGACACTGCCAACTATACCTGCGTG 372
Db
        667 GCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAAC 726
Qу
           Db
        373 GCCAAGAACATCGTGGCCAAACGCCGGAGCACCACTGCCACCGTCATCGTCTACGTGAAT 432
        727 GGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGG 786
Qy
                                   433 GGCGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGCTCCAACCGCTGTGGCCGAGGCTGG 492
Db
        787 CAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAG 846
Qy
           493 CAGAAGCGCACCCGGACCTGCACCAACCCCGCTCCACTCAACGGAGGGGCCTTCTGCGAG 552
Db
```

847 GGGCAGAATGTCCAGAAAACAGCCTGCGCCACCTGTGCCCAGTAGACGGCAGCTGGAGC 906

Adams, M.D. and Cargill, M.

Direct Submission

TITLE

Qy

11 111

Db	553	GGCCAGGCATTCCAGAAGACCGCCTGCACCACCATCTGCCCAGTCGATGGGGCGTGGACG	612
Qy	907	CCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGC	966
Db	613	GAGTGGAGCAAGTGGTCAGCCTGCAGCACTGAGTGTGCCCACTGGCGTAGCCGCGAGTGC	672
Qу	.967	TCTGACCCAGCACCCGCAACGGAGGGGAGGGGGGGGGGG	1026
Db	673	ATGGCGCCCCACCCCAGAACGGAGGCCGTGACTGCAGCGGGACGCTGCTCGACTCTAAG	732
Qу	1027	AACTGTACCAGTGACCTCTGTGTACACAGTGCTTCT	1062
Db	733	AACTGCACAGATGGGCTGTGCATGCAAAATAAGAAAACTCTAAGCGACCCCAACAGCCAC	792
Qy .	1063	GGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTC	1110
Db	793	CTGCTGGAGGCCTCAGGGGATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCATCTTCGTG	852
Qу	1111	TGCCTGGTCCTGCTGCTGCTCTCTCTCTCTTTATTGCCGGAAGAAGGAGGGGCTG	1170
Db	853	GTCGTGGCAATCCTCATGGCGGTGGGGGTGGTGTGTACCGCCGCAACTGCCGTGACTTC	912
Qу	1171	GACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATC	1227
Db	913	GACACAGACATCACTGACTCTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAACTTT	972
Qу	1228	AAGCCCAGCAAAGCAGACCCCCATCTGCTCACCATCCAGCCGGACCTCAGC	1281
Db	973	AAGACGGCAAGCCCAGCACCCGCAGCTCCTACACCCCTCTGTGCCTCCTGACCTGACA	1032
Qу	1282	ACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGG	1329
Db		GCCAGCGCCGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCCACCGACAAAATC	
QУ	1330	CCCAGCCCCAAGTTCCAGCTCACCA	1354
Db	1093	CCCATGACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGC	1152
Qу	1355	ATGGGCACCTGCTCAGCCCCCTGGGTGGCGCCCCCACACA	1395
Db	1153	TCCAGCACCACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGGGTCTTG	1212
Qу	1396	CTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAG	1455
Db	1213	CCGCCTGGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGCGCAGC	1272
Qу	1456	AACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1500
Db	1273	GCCAGCCTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGGGAGCAGCGTCAGC	1332
Qу	1501	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1560
Db	1333	GGCACCTTTGGCTGCCTGGGTGGGAGGCTCAGCATCCCCGGCACAGGGGTCAGCTTGCTG	1392
Qу	1561	ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Db	1393	GTGCCCAATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAAG	1452

Qy ·	1621	CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC	1680
Db	1453	GCAGAAAGTACCCTCCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCGGTGACC	1512
Qу	1681	TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1740
Db	1513	TGTGGACCCACAGGCCTCCTGCTGTGCCGCCCCGTCATCCTCACCATGCCCCACTGTGCC	1572
Qу	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Db	1573	GAAGTCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCACTGG	1632
Qy .	1801	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG	1860
Db	1633	GAGGAGGTGGTGACCCTGGATGAGGAGACCCTGAACACCCCTGCTACTGCCAGCTGGAG	1692
Qу	1861	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	1920
Db	1693	CCCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCC	1752
Qу	1921	CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	1980
Db	1753	TATTCCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCCGTCTTCGCCCCCCCC	1812
Qу	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2040
Db	1813	TCCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAG	1872
Qу	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC	2100
Db	1873	GTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATG	1932
Qу	2101	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG	2160
Db	1933	TTCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCATGCCCATTGG	1992
Qу	2161	AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG	2220
Db	1993	AGGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGC	2052
Qу	2221	CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG	2280
Db	2053	CAGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTC	2112
Qу	2281	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2340
Db	2113	ACCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACC	2172
Qу	2341	AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGG	2397
Db	2173	ACTCTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACT	2232
Qу	2398	GTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATA	2457
Db	2233	GTCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATA	2292

```
Qy
        2458 ATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAA 2517
                  11111
Db
        2293 TGCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAG 2352
        2518 CTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATC 2577
Qу
                   2353 CTCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTGATC 2412
Db
        Qу
                 1 1 11 111111
                                                          2413 CTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGTGCC 2472
Dh
        2638 GTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
Qу
                                                     1 11 1 11 1111
                      1111111
                                                  1
              -11
                                          1
        2473 TTGGAGGAGATGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGCTGA 2532
Db
RESULT 8
AY411749
LOCUS
           AY411749
                                  2532 bp
                                            DNA
                                                    linear
                                                            GSS 12-DEC-2003
          Mus musculus HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION
           genomic survey sequence.
ACCESSION
           AY411749
           AY411749.1 GI:39767717
VERSION
KEYWORDS
           GSS.
           Mus musculus (house mouse)
SOURCE
  ORGANISM Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
              (bases 1 to 2532)
REFERENCE
           Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
  AUTHORS
           Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
           Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
           Adams, M.D. and Cargill, M.
  TITLE
           Inferring nonneutral evolution from human-chimp-mouse orthologous
           gene trios
  JOURNAL
           Science 302 (5652), 1960-1963 (2003)
  PUBMED
           14671302
REFERENCE
             (bases 1 to 2532)
  AUTHORS
           Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
           Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
           Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
           Adams, M.D. and Cargill, M.
           Direct Submission
  TITLE
           Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
  JOURNAL
           Rockville, MD 20850, USA
           This sequence was made by sequencing genomic exons and ordering
COMMENT
           them based on alignment.
                   Location/Qualifiers
FEATURES
                   1. .2532
    source
                   /organism="Mus musculus"
                   /mol type="genomic DNA"
                   /db xref="taxon:10090"
                   <1. .>2532
    gene
                   /locus tag="HCM4327"
ORIGIN
```

Query M		28.9%; Score 780.4; DB 9; Length 2532;	
		Similarity 60.1%; Pred. No. 5.1e-156; 5; Conservative 0; Mismatches 876; Indels 129; Gaps	8;
Qy	307	GAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATAC	366
Db			
Qy		TGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATC	
Db		TGGTGCCAGTGCGTGGCCTGGAGCTCTTCGGGAACTACCAAGAGTCGCCGAGCCTACATC	
١.		·	
Qy 		CGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTG	
Db		CGCATTGCCTACTTGCGCAAGAACTTTGACCAGGAGCCTCTGGCCAAGGAGGTACCCTTG	
Qу		GAGCAGGCATCGTGCTGCCCTGCCGTCCACCGGAGGCATCCCTCCAGCCGAGGTGGAG	
Db	193	GATCATGAGGTCCTTCTGCAGTGCCGCCCACCGGAGGGAG	252
Qу	547	TGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGG	606
Db	253	TGGCTCAAGAATGAAGATGTCATTGACCCCGCTCAGGACACTAACTTCCTGCTCACCATT	312
QУ	607	GAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTG	666
Db	313	GACCACAACCTCATCCGCCAGGCGCCCTCTCAGACACGGCCAACTACACCTGTGTG	372
Qу	667	GCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAAC	726
Db	373	GCCAAGAATATCGTGGCCAAGCGCCGGAGCACCACGGCCACAGTCATCGTCTATGTGAAT	432
Qу	727	GGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGG	786
Db	433		492
Qy	787	CAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAG	846
Db	493		552
Qу		GGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGC	906
Db			612
Qу	907	CCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGC	966
Db	613		672
Qу	967	TCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGC	1026
Db	673	ATGGCACCGCCACCCCAGAACGGAGGCCGTGACTGCAGCGGGACGCTACTTGACTCCAAG	732
Qy :	1027	AACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGAC	1074
Db			
		GTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGC	

Db	793	CCCCTGGAGACATCGGGAGATGTGGCACTGTACGCAGGCCTTGTGGTGGCCGTCTTTGTG	852
Qy Db		CTGGTCCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTG	
Qу	1171	GACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCGTCAGCATC	1227
Db	913	GACACGGACATCACCGACTCCTCTGCGGCCCTCACTGGTGGCTTCCACCCTGTCAACTTC	972
Qу	1228	AAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGC	1281
Db	973	AAGACTGCAAGGCCCAACAACCCGCAGCTCCTGCACCCGTCCGCCCCTCCAGACCTAACG	1032
Qу	1282	ACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGG	1328
Db	1033	GCCAGTGCTGGCATCTACCGCGGGCCTGTGTATGCCCTGCAGGACTCCGCCGACAAGATC	1092
Qу	1329	GCCCAGCCCCAAGTTCCAGCTCACCAAT	1356
Db	1093	CCCATGACTAATTCGCCCCTGCTGGATCCCCTGCCCAGCCTCAAGATCAAGGTCTATAAC	1152
Qу	1357	GGGCACCTGCTCAGCCCCCTGGGTGGCGGC	1386
Db	1153		1212
Qу	1387	CGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTC	1446
Db	1213	CCGCCGGGCACGTACCCAGGCGATTTCTCCCGGGACACCCATTTCCTGCACCTGCGCAGT	1272
Qу	1447	TCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1500
Db	1273	GCCAGCCTTGGTTCCCAGCACCTCCTGGGCCTACCTCGGGACCCCAGCAGCAGTGTCAGC	1332
Qу	1501	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1560
Db	1333	GGCACCTTTGGTTGCCTGGGAGGAAGGCTGAGCCTCCCCGGCACAGGGGTCAGCCTGTTG	1392
QУ	1561	ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Db	1393	GTACCAAATGGAGCCATTCCCCAGGGCAAGTTCTATGACCTGTATCTACATATCAACAAG	1452
Qу	1621	CCGGAAGACGTGAGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC	1680
Db	1453	GCCGAAAGCACCCTCCCACTTTCAGAAGGTTCCCAGACAGTATTGAGCCCCTCGGTGACC	1512
Qу	1681	TGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1740
Db	1513	TGTGGGCCCACAGGCCTACTCCTGTGCCGCCCTGTCGTCCTCACCGTGCCCCACTGTGCT	1572
Qу	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Db	1573	GAAGTCATCGCTGGAGACTGGATCTTTCAGCTCAAGACCCAGGCCCATCAGGGCCACTGG	1.632
Qy .	1801	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG	1860

	Db	1633	GAGGAGGTGGTGACCTTGGATGAGGAGACCCTCAACACCCCTGCTACTGCCAGCTGGAG	1692
	Qy	1861	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	1920
•	Db	1693	GCTAAGTCCTGCCACATCCTGGACCAGCTGGGTACCTACGTATTCATGGGCGAGTCC	1752
	Qу	1921	CTCAGCGTGGCTGCCCCAAGCCTCCTCTGTTTGCGCCGGTGGCCTGCACC	1980
	Db	1753	TACTCTCGCTCTGCAGTCAAGCGGCTCCAGCTGGCCATCTTCGCCCCAGCCCTCTGCACC	1812
	Qу	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2040
	Db	1813	TCCCTGGAGTATAGCCTCAGGGTCTACTGTCTGGAGGACACCCTGTAGCACTGAAGGAG	1872
	Qy	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC	2100
	Db	1873	GTCCTGGAGCTGGAGAGCCTCTGGGTGGCTACTTGGTGGAGGAGCCCAAGCCTTTGCTC	1932
	Qу	2101	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG	2160
	Db	1933	TTTAAGGACAGTTACCACAACCTACGCCTCTCCCTCCATGACATCCCCCATGCCCACTGG	1992
	Qу	2161	AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG	2220
	Db	1993	AGGAGCAAACTACTGGCCAAGTACCAGGAGATTCCCTTCTACCACGTCTGGAATGGCAGC	2052
	QУ	2221	CAGCGGTACTTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG	2280
	Db	2053	CAGAGAGCCCTGCACTGCACTTCACCCTGGAGAGGCATAGCCTGGCCTCCACGGAGTTC	2112
	Qу	2281	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2340
	Db	2113	ACCTGTAAGGTCTGCGGCAGGTCGAAGGGGAAGGCCAGATTTTCCAGCTGCACACA	2172
	Qy	2341	AACATCACCAAGGACACAAGGTTTGCTGAGCTGGCTCTGGAGAGTGAAGCGGGG	2397
	Db	2173	ACGTTGGCCGAGACGCCTGCTGGCTCCCTGGATGCTCTCTGCTCTGCCCCGGGCAATGCC	2232
	QУ	2398	GTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATA	2457
	Db	2233	ATCACCACCCAGCTGGGACCCTATGCCTTCAAGATACCCCTGTCCATCCGCCAAAAGATC	2292
	Qу	2458	ATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAA	2517
	Db	2293	TGCAGCAGCCTGGACGCCCCAACTCCCGGGGCAACGACTGGAGGCTGTTGGCGCAGAAG	2352
	Qу	2518	CTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATC	2577
	Db	2353	CTGTCCATGGACCGGTACCTAAACTACTTCGCCACCAAAGCTAGTCCCACAGGTGTCATC	2412
	Qу	2578	CTCAACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC	2637
	Db	2413	TTAGACCTCTGGGAAGCTCGGCAACAGGATGACGGGGACCTCAACAGCCTGGCCAGTGCC	2472
•	Qу	2638	GTGGCTGGACTGGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA	2697
	Db	2473	TTGGAGGAGATGGCAAGAGTGAGATGCTGGTAGCCATGGCCACAGATGGCGATTGCTGA	2532

•

```
BI758231
                                                              EST 25-SEP-2001
                                   1034 bp
                                              mRNA
                                                      linear
LOCUS
           BI758231
           603029876F1 NIH MGC 114 Homo sapiens cDNA clone IMAGE:5200171 5',
DEFINITION
           mRNA sequence.
ACCESSION
           BI758231
           BI758231.1 GI:15749809
VERSION
KEYWORDS
           EST.
           Homo sapiens (human)
SOURCE
  ORGANISM
           Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
           1 (bases 1 to 1034)
REFERENCE
           NIH-MGC http://mgc.nci.nih.gov/.
  AUTHORS
           National Institutes of Health, Mammalian Gene Collection (MGC)
  TITLE
           Unpublished (1999)
  JOURNAL
           Contact: Robert Strausberg, Ph.D.
COMMENT
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM11501 row: q column: 20
           High quality sequence stop: 793.
FEATURES
                    Location/Qualifiers
                    1. .1034
     source
                    /organism="Homo sapiens"
                    /mol type="mRNA"
                    /db xref="taxon:9606"
                    /clone="IMAGE: 5200171"
                    /lab host="DH10B"
                    /clone lib="NIH MGC 114"
                    /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
                    Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
                    male brains, age range 23-27 yo. Library is oligo-dT
                    primed and directionally cloned (EcoRV site is destroyed
                    upon cloning). Average insert size 1.5 kb, insert size
                    range 1-3 kb. Library is normalized and enriched for
                    full-length clones and was constructed by C. Gruber
                    (Invitrogen). Research Genetics tracking code 019. Note:
                    this is a NIH MGC Library."
ORIGIN
                                 Score 768.6; DB 4;
                                                      Length 1034;
                         28.5%;
                         91.2%; Pred. No. 1.5e-153;
  Best Local Similarity
  Matches 918; Conservative
                                0; Mismatches
                                                 74;
                                                      Indels
                                                              15;
                                                                   Gaps
                                                                            9;
          105 GCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACAT 164
Qу
              1 GCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACAT 60
Db
         165 CGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTT 224
Qy
```

RESULT 9

	Db	61	CGTCAAGAACAAGCCAGTGCTGCTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTT	120
	Qу	225	CAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGG	284
	Db	121	CAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGG	180
	Qу	285	GAGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAA	344
	Db	181	GAGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAA	240
	Qу	345	GGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCAC	404
	Db	241	GGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCAC	300
	Qу	405	CAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCC	464
	Db	301	CAAGAGTCAGAAGGCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCC	360
	Qу	465	GCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGG	524
	Db	361	GCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGG	420
	QУ	525	CATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGA	584
	Db	421	CATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGA	480
	Qу	585	CCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGA	644
	Db	481	CCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGA	540
	Qу	645	CACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGC	704
	Db	541	${\tt CACGGACAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGC}$	600
	Qу	705	TGCTGTCATCGTCAACGGTGGGTGGTCGA-CGTGGACCGAGTGGTCCGTCTGCA	763
	Db	601	TGCTGTCATCGTCAACGTGAACGGTGGTCGACCGTGGACCGAGTGGTCCGTCTGC-	659
	Qy	764	GCGCCAGCTGTGGGCGCGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTC	823
	Db	660	GCGCCAGCTGTGGGCGGCTGGCAGAAACCGGAGCCGGAGCTGCACAACCCGGTGCCTC	719
	Qу	824	TCAACGGGGGCGCTTTCTGTGA-GGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTG	882
,	Db	720	TCAACGGGGGCGCTTTCTGTGAGGGGGCAGAATGTCCAGAAAGCAGCTGCGCCACCT	776
	QУ	883	TGCCCAGTAGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACT	940
	Db	777	GTGCCAGTGGACGCAGCTGTAGCACGTGGAGCCAGTGGTCGGCCTGTTGGGCTT	836
	Qу	941	GCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAG	998
	Db	8.3.7	GCACCCACTTGGGCGGAGCCGGAGTGCTCTGAACCCAGCACCCCGGCACGGCAGGGGGAG	896
	Qу	999	GTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAG	1055

```
Db
          897 GTGTGCCCAGGTCACTGGACCTGGCACCCGGGA-TGGTCCAGTGAGCTCTGTGT-CCCAC 954
         1056 TGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCG 1102
Qy
                955 GGGTTCTGGCCCTGAGGACTTGGCCTCCTATGTGGGCCTCATCCCGG 1001
Db
RESULT 10
BC033727
LOCUS
            BC033727
                                     1532 bp
                                                mRNA
                                                        linear
                                                                  HTC 25-MAR-2004
            Homo sapiens unc-5 homolog A (C. elegans), mRNA (cDNA clone
DEFINITION
            IMAGE:5166762), containing frame-shift errors.
ACCESSION
            BC033727.1 GI:21707230
VERSION
            HTC.
KEYWORDS
            Homo sapiens (human)
SOURCE
  ORGANISM
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
               (bases 1 to 1532)
  AUTHORS
            Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
            Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
            Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
            Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
            Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
            Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
            Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
            Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
            Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
            McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
            Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
            Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
            Fahey, J., Helton, E., Ketteman, M., Madan, A., Rodrigues, S.,
            Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
            Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
            Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
            Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
            Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
  TITLE
            Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL
            Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
            12477932
   PUBMED
               (bases 1 to 1532)
REFERENCE
  AUTHORS
            Strausberg, R.
            Direct Submission
  TTTLE
            Submitted (02-JUL-2002) National Institutes of Health, Mammalian
  JOURNAL
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            NIH-MGC Project URL: http://mgc.nci.nih.gov
  REMARK
COMMENT
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
```

Sequencing Center (NISC),

```
Gaithersburg, Maryland;
          Web site: http://www.nisc.nih.gov/
          Contact: nisc mgc@nhgri.nih.gov
          Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
          Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
          Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
          Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
          Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
          McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
          Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
          Young, A., Zhang, L.-H. and Green, E.D.
          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
          Series: IRAK Plate: 68 Row: i Column: 2
          This clone was selected for full length sequencing because it
          passed the following selection criteria: GenomeScan gene prediction
          This clone has the following problem: frame shifted.
                  Location/Oualifiers
FEATURES
                  1. .1532
    source
                  /organism="Homo sapiens"
                  /mol type="mRNA"
                  /db xref="taxon:9606"
                  /clone="IMAGE:5166762"
                  /tissue type="Brain, adult medulla"
                  /clone \overline{lib}="NIH MGC 119"
                  /lab host="DH10B"
                  /note="Vector: pCMV-SPORT6"
ORIGIN
 Query Match
                       27.3%;
                             Score 736.6; DB 3;
                                                Length 1532;
 Best Local Similarity
                            Pred. No. 1.1e-146;
                      84.2%;
 Matches 917; Conservative
                             0; Mismatches 4;
                                                       168;
                                                                    1;
                                                Indels
                                                             Gaps
        292 GGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTC 351
Qу
            612 GGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTC 671
Db
        352 GGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGT 411
Qу
            672 GGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGT 731
Db
        412 CAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCC 471
Qy
            732 CAGAAGGCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCC 791
Db
Qу
        472 AAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCT 531
            792 AAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCTGCCGTCCACCGGAGGGCATCCCT 851
Db
        532 CCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAAT 591
Qу
            852 CCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAAT 911
Db
        592 GTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCC 651
Qу
```

912 GTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCC 971

Db

Qу	652	AACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTC	711
Db	972	AACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTTC	1031
Qу	712	ATCGTCTACGTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGC	771
Db	1032	ATCGTCTACGTG	1043
Qу	772	TGTGGGCGCGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGG	831
Db .	1044		1043
Qу	832	GGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTA	891
Db	1044		1043
Qу	892	GACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGG	951
Db	1044	GACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGG	1103
Qу	952	CGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAG	1011
Db	1104		1163
Qу	1012	GACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAG	1071
Db	1164	GACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAG	1223
Qy	1072	GACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC	1131
Db	1224		1283
Qу	1132	GTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGGCTGGACTCAGATGTGGCTGACTCG	1191
Db	1284	GTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCG	1343
Qу	1192	TCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCC	1251
Db	1344	TCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCC	1403
Qу	1252	CATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCAGGGCAGTCTC	1311
Db	1404	CATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCAGGGCAGTCTC	1463
Qy.	1312	TGTCCCCGGCAGGATGGGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGC	1371
Db	1464	TGTCCCCGGCAGGATGGGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGC	1523
Qу	1372	CCCCTGGGT 1380	
Db ·	1524	CCCCTGGGT 1532	

RESULT 11 AI951556

```
wv36f04.xl NCI CGAP Ov18 Homo sapiens cDNA clone IMAGE:2531647 3'
DEFINITION
           similar to TR:008721 008721 TRANSMEMBRANE RECEPTOR UNC5H1. ;, mRNA
           sequence.
ACCESSION
           AI951556
VERSION
           AI951556.1 GI:5743866
KEYWORDS
           EST.
SOURCE
           Homo sapiens (human)
  ORGANISM
           Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
              (bases 1 to 788)
  AUTHORS
           NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE
           National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
           Tumor Gene Index
  JOURNAL
           Unpublished (1997)
           Contact: Robert Strausberg, Ph.D.
COMMENT
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
           R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento
           Soares, Ph.D. cDNA Library Arrayed by: Christa Prange, The
           I.M.A.G.E. Consortium DNA Sequencing by: Washington University
           Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           www-bio.llnl.gov/bbrp/image/image.html
           Insert Length: 1125
                               Std Error: 0.00
           Seq primer: -40UP from Gibco
           High quality sequence stop: 446.
FEATURES
                    Location/Qualifiers
                    1. .788
    source
                    /organism="Homo sapiens"
                    /mol type="mRNA"
                    /db xref="taxon:9606"
                    /clone="IMAGE:2531647"
                    /tissue type="fibrotheoma"
                    /lab host="DH10B (phage-resistant)"
                    /clone lib="NCI CGAP Ov18"
                    /note="Organ: ovary; Vector: pT7T3D-Pac (Pharmacia) with a
                    modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
                    strand cDNA was primed with a Not I - oligo(dT) primer [5'
                    double-stranded cDNA was ligated to Eco RI adaptors
                    (Pharmacia), digested with Not I and cloned into the Not
                    I and Eco RI sites of the modified pT7T3 vector. Library
                    went through one round of normalization, and was
                    constructed by Bento Soares and M. Fatima Bonaldo.
ORIGIN
 Query Match
                        27.3%; Score 735.2; DB 1;
                                                    Length 788;
 Best Local Similarity
                        96.5%; Pred. No. 2e-146;
 Matches 749; Conservative
                               0; Mismatches
                                                    Indels
                                                              0;
                                                                  Gaps
                                               27;
        1424 CCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCA 1483
Qу
             Db
           1 CCGAGGAGTTCGTCTCCGCCTCTCCGCCCAGAACTACTTCCGCTCCCTGCCCCGAGGCA 60
Qу
        1484 CCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATA 1543
```

Db	61	
Qу	1544	CAGGTATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCT 1603
Db	121	CAGGAATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCT 180
Qу	1604	ACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTG
Db	181	ACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTG
Qу	1664	TGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGG 1723
Db	241	TGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTTCTGCTCACCCGGCCAGTCATCCTGG 300
Qу	1724	CTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGT 1783
Db	301	CTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGT 360
Qу	1784	CGTGCGAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCT 1843
Db	361	CGTGCGAGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCT 420
QУ	1844	ACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTG 1903
Db	421	ACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGAGCCGCTATG 480
Qу	1904	CCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTG 1963
Db	481	CCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTG 540
Qу	1964	CGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC
Db	541	CGCCGGTGGCCTCCCTCGAGTACAACATACTGGTCTACTGCCTGC
Qу	2024	ACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGG 2083
Db	601	ACGATGCACTCAACGTAGTGGTGCAGCTGGAGAAGCAGCTGCAGGGACAGCTGATCCAGG 660
Qу	. 2084	AGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC
Db	661	AGCCACTGGTACTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATNCATCCACGATG 720
Qу	2144	TGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTT 2199
Db	721	TGCCCAGCTNCCNTGTGGAGAGTAAGCTTCTTGTCAGCTACCCAGAGATCCNCTAT 776

RESULT 12 BX348193/c

LOCUS BX348193 796 bp mRNA linear EST 08-APR-2004 DEFINITION BX348193 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens

cDNA clone CS0DB008YE02 5-PRIME, mRNA sequence.

ACCESSION BX348193

VERSION BX348193.2 GI:46286231

KEYWORDS EST.

SOURCE Homo sapiens (human)

```
ORGANISM
          Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
             (bases 1 to 796)
           Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 AUTHORS
           Full-length cDNA libraries and normalization
 TITLE
 JOURNAL
           Unpublished (2001)
           On May 5, 2003 this sequence version replaced gi:30367258.
COMMENT
           Contact: Genoscope
           Genoscope - Centre National de Sequencage
           2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
           Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
           1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
           end enriched, double-strand cDNA was digested with Not I and cloned
           into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
           was normalized. Library was constructed by Life Technologies, a
           division of Invitrogen. This sequence belongs to sequence cluster
           3239.r
           For more information about this cluster, see
           http://www.genoscope.cns.fr/cdna?s=CS0BAF004ZD01 AF00293 1&c=3239.r
                   Location/Qualifiers
FEATURES
                   1. .796
    source
                   /organism="Homo sapiens"
                   /mol type="mRNA"
                   /db xref="taxon:9606"
                   /clone="CS0DB008YE02"
                   /tissue type="NEUROBLASTOMA COT 10-NORMALIZED"
                   /clone lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
                   /note="1st strand cDNA was primed with a NotI-oligo(dT)
                   primer. Five prime end enriched, double-strand cDNA was
                   digested with Not I and cloned into the Not I and EcoR V
                   sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
                               Score 721.6; DB 5;
 Query Match
                       26.8%;
                                                  Length 796;
 Best Local Similarity
                       96.4%; Pred. No. 1.7e-143;
                                                                      2;
 Matches 758; Conservative
                              0; Mismatches
                                             26;
                                                  Indels
                                                           2;
                                                               Gaps
        1634 GGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTG 1693
Òу
             795 GTTGCCCCAAACCGGGCTGTCAGACCCTGTTGAGTCCCATGGTTAGCTGTGAA-CCCCTG 737
Db
        1694 GCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTG 1753
Qу
             736 GCGTCCTGCTCACCGGCCCAGTCATCCTGGGTATGGACCACTGTGGGGAGCCCAGCCCTG 677
Db
        1754 ACAGCTGGAGCCTGCGCCTC-AAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCTG 1812
Qу
                    676 ACAGCTGAGGCCTGCGCCTCAAAAAAACAGTCGTGCGAGGACAGCTGGGAGTATGTCCTG 617
Db
        1813 CACCTGGGCGAGGGGGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGC 1872
Qу
             Db
         616 CACCTGGGCGAGNAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGC 557
```

1873 TACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCT 1932

Qу

Db	556	TACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCT	497							
Qy	1933	GCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC	1992							
Db	496	CCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC								
Qу	1993	AACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTG	2052							
Db	436	AACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTG	377							
Qy	2053	GAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGT	2112							
Db	376	GAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGT	317							
Qy	2113	TACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTC	2172							
Db	316	TACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTC	257							
Qy	2173	CTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTG	2232							
Db	256	CTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTG	197							
Qу	2233	CACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTG	2292							
Db	196	CACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTG	137							
Qу	2293	TGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACATCACCAAG	2352							
Db	136	TGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACATCACCAAG	77							
Qу	2353	GACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTG	2412							
Db	76	GACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGNCCCTGGT	17							
Qу	2413	GGCCCC 2418								
Db	16	GGGCCC 11								
RESULT 13	3									
BI818609 LOCUS	R	I818609 818 bp mRNA linear EST 04-OCT-	-2001							
DEFINITIO	ом 6	03033362F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5174559 S								
ACCESSION	N B	1818609								
VERSION		I818609.1 GI:15929902 ST.								
KEYWORDS SOURCE	_	omo sapiens (human)								
ORGANIS	SM H	omo sapiens								
		ukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostor ammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	ni;							
REFERENCI										
AUTHOR		<pre>IH-MGC http://mgc.nci.nih.gov/.</pre>								
TITLE JOURNA		National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)								
COMMENT		ontact: Robert Strausberg, Ph.D.								
		mail: cgapbs-r@mail.nih.gov								

```
Tissue Procurement: Life Technologies, Inc.
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM11434 row: 1 column: 16
          High quality sequence stop: 744.
                  Location/Qualifiers
FEATURES
                  1. .818
    source
                  /organism="Homo sapiens"
                  /mol type="mRNA"
                  /db xref="taxon:9606"
                  /clone="IMAGE:5174559"
                  /lab host="DH10B"
                  /clone lib="NIH MGC 115"
                  /note="Organ: pooled brain, lung, testis; Vector:
                  pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
                  source anonymous pool of 6 male brains, age range 23-27; 1
                  male lung, age 27; and 1 male testis, age 69. Library is
                  oligo-dT primed and directionally cloned (EcoRV site is
                  destroyed upon cloning). Average insert size 1.8 kb,
                  insert size range 1-3 kb. Library is normalized and
                  enriched for full-length clones and was constructed by C.
                  Gruber (Invitrogen). Research Genetics tracking code
                  021. Note: this is a NIH MGC Library."
ORIGIN
                      26.5%;
 Query Match
                             Score 713.4; DB 4;
                                               Length 818;
 Best Local Similarity
                      97.6%;
                             Pred. No. 9.4e-142;
 Matches 745; Conservative
                            0; Mismatches
                                                                   2;
                                           16;
                                                Indels
                                                         2;
          1 ATGGCCGTCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
Qy
            4 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 63
Db
         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
Qy
            64 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCGGTGCCTGGTGCCAACCCG 123
Db
        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
Qv
            111411141141414
        124 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 183
Db
        181 GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
Qy
            Db
        184 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 243
        241 TGGGTGCGCCAGGTGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
Qу
            244 TGGGTGCGCCAGGTGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 303
Db
```

301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360

304 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 363

Qy'

Db

Qy	361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
Db	364 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 423
Qу	421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
Db	424 TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 483.
Qy	481 TCCCTGGAGCAGGGCATCGTGCCGTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
Db	484 TCCCTGGAGCAGGGCATCGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 543
Qy	541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
Db	544 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 603
Qy	601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
Db	
Qy	661 TGCGTGGCCAAGAACATCGTGGCACGTC-GCCGCAGCGCCTCCGCTGCTGTCATCGTCTA 719
Db	664 TGCGTGGCCAAGAACATCGTGGCACGTCAGCCGCAGGGCCTCCGCTGCTGTCATCGTCTA 723
Qу	720 CGTGAACGGTGGTCGACG-TGGACCGAGTGGTCCGTCTG 761
Db	724 CGTGGACGCAGCTGGAGCCCGTTGGAGCCAGTGGTCGGGCTG 766
RESULT 14 AY411748 LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM	AY411748 2532 bp DNA linear GSS 12-DEC-2003 Pan troglodytes HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence. AY411748 AY411748.1 GI:39767716 GSS. Pan troglodytes (chimpanzee) Pan troglodytes Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
AUTHORS	•
m T m T T	
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering

them based on alignment. FEATURES Location/Qualifiers

source 1. .2532

/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"

gene <1. .>2532

/locus_tag="HCM4327"

ORIGIN

		Similarity 54.4%; Pred. No. 3.7e-134;	
	Matches 137	0; Conservative 0; Mismatches 1021; Indels 129; Gaps	9;
Qу	307	GAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATAC	366
Db	13	GAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTCTTTGGGCTGGAGGATTAC	72
Qy	367	TGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATC	426
Db	73	TGGTGCCAGTGCGTGGCCTGGAGCTCTGCGGGCACCACCAAGAGTCGCCGAGCCTACGTC	132
Qу	427	CGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTG	486
Db	133	CGCATCGCCTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	192
Qу	487	GAGCAGGCATCGTGCCGTGCCGTCCACCGGAGGCATCCCTCCAGCCGAGGTGGAG	546
Db	193	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	252
Qу	547	TGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGG	606
Db	253	TGGCTCAAGAATGAGGATGTCATCGACCCCACCCAGGACACCAACTTCCTGCTCACCATC	312
Qу	607	GAGCACAGCCTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTG	666
Db	313	GACCACAACCTCATCCTCCGCCAGGCCCGCCTGTCGGACACTGCCAACTATACCTGCGTG	372
Qу	667	GCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAAC	726
Db	373	GCCAAGAACATCGTGGCCAAACGCCGGAGCACCACTGCCACCGTCATCGTCTACGTGAAT	432
Qу	727	GGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCGGCTGG	786
Db	433	GGCGGCTGGTCCAGCTGGGCAGAGTGGTCGCCCTGCTCCAACCGCTGTGGNNGANNCTGG	492
Qу	787	CAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAG	846
Db	493	CAGAAGCGCACCCGGACCTGCACCAACCCCGCCCCACTCAACGGAGGGNNNNTCTGCGAG	552
Qу	847	GGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGC	906
Db	553	GGCCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	612
Qу	907	$\tt CCGTGGAGCAAGTGGTCGGCCTGTGGGCTGACTGCACCCACTGGCGGAGCCGTGAGTGC$	966

Db	613		672
Qу	967	TCTGACCCAGCACCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGC	1026
Db	673	ATGGCGCCCCACCCCAGAACGGAGGCCGTGACTGCAGCGGGACGCTGCTCGACTCTAAG	732
Qу	1027	AACTGTACCAGTGACCTCTGTGTACACAGTGCTTCT	1062
Db	733		792
Qу	1063	GGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTC	1110
Db	793		852
Qу	1111	TGCCTGGTCCTGCTGCTGCTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTG	1170
Db	853	GTCNTGGCAATCCTCATGGCGGTGGGGGTGGTGTACCGCCGCAACTGCCGTGACTTC	912
Qу	1171	GACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATC	1227
Db	913	GACACAGACATCACTGACTCTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAACTTT	972
Qу	1228	AAGCCCAGCAAAGCAGACCCCCATCTGCTCACCATCCAGCCGGACCTCAGC	1281
Db	973	AAGACGGCAAGGCCCAGCAACCCGCAGCTCCTACACCCCTCTGTGCCTCCTGACCTGACA	1032
Qу	1282	ACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGG	1329
Db	1033		1092
Qу	1330	CCCAGCCCCAAGTTCCAGCTCACCA	1354
Db	1093	CCCATGACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGC	1152
Qy	1355	ATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACA	1395
Db	1153	TCCAGCACTACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGAGTCTTG	1212
Qy ·	1396	CTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAG	1455
Db	1213	CCGCCTGGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGCGCAGC	1272
Qу	1456	AACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1500
Db	1273	GCCAGCCTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGNNAGCAGCGTCAGC	1332
Qy	1501	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1560
Db	1333		1392
Qу	1561	ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Db	1393		1452
Qу	1621	CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC	1680

Db	1453	GCAGAAAGTACCCTCCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCAGTGACC	1512
Qy .	1681	TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1740
Db	1513	TGTGGACCCACAGGCCTCCTGCTGTGCCACCCCGTCATCCTCACCATGCCCCACTGTGCC	1572
Qу	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Db	1573	GAAGTCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCACTGG	1632
Qу	1801	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG	1860
Db	1633	GAGGAGGTGGTGACCCTGGATGAGGAGACCCTGAACACCCCTGCTACTGCCAGCTGGAG	1692
Qу	1861	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	1920
Db	1693	CCCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCC	1752
Qy	1921	CTCAGCGTGGCTGCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	1980
Db	1753	TATTCCCNNNNNCAGTCNNNNNGCTCCNGCTGGCCGTCTTNGCCCCTGCCCCTCTGCACC	1812
Qy	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2040
Db	1813	TCCCTGGANNNNNGCCTCCGGGTCTACTGCCTGGNNNNNNNNNN	1872
Qy ·	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC	2100
Db	1873	GTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATG	1932 .
Qу	2101	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG	2160
Db	1933	TTCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGG	1992
Qу	2161	AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG	2220
Db	1993	AGGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTGTCACATTTGGAGTGGCAGC	2052
Qу	2221	CAGCGGTACTTGCACCTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG	2280
Db	2053	CAGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTC	2112
Qy	2281	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2340
Db	2113	ACCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACC	2172
Qу	2341	AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGG	2397
Db	2173	ACTCTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACT	2232
Qy	2398	GTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATA	2457
Db	2233	GTCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATA	2292
Qy	2458	ATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAA	2517
Db	2293	TGCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAG	2352

```
2518 CTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATC 2577
Qу
             \Pi
                   Db
        2353 CTCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTGATC 2412
        Qy
               11111
        2413 CTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGTGCC 2472
Db
        2638 GTGGCTGGACTGGGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
Qy
                     - 11
                                                 1
                                                     1 11 1 11 11111
              111.
                                          1
        2473 TTGGAGGAGATGGCAAGAGTGAGATGCTGGTGGCCTGTGGCCACCGACGGGGACTGCTGA 2532
Dh
RESULT 15
BF311804
                                                           EST 21-NOV-2000
LOCUS
           BF311804
                                  934 bp
                                           mRNA
                                                   linear
           601897316F1 NIH MGC 19 Homo sapiens cDNA clone IMAGE:4126706 5',
DEFINITION
           mRNA sequence.
           BF311804
ACCESSION
           BF311804.1 GI:11259566
VERSION
KEYWORDS
SOURCE
           Homo sapiens (human)
 ORGANISM
           Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
              (bases 1 to 934)
 AUTHORS
           NIH-MGC http://mgc.nci.nih.gov/.
           National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE
 JOURNAL
           Unpublished (1999)
           Contact: Robert Strausberg, Ph.D.
COMMENT
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
           Plate: LLCM1016 row: p column: 03
           High quality sequence stop: 707.
FEATURES
                   Location/Qualifiers
                   1. .934
    source
                   /organism="Homo sapiens"
                   /mol type="mRNA"
                   /db xref="taxon:9606"
                   /clone="IMAGE:4126706"
                   /tissue type="neuroblastoma"
                   /lab host="DH10B (phage-resistant)"
                   /clone lib="NIH MGC 19"
                   /note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
                   EcoRI; cDNA made by oligo-dT priming. Directionally
                   cloned into EcoRI/XhoI sites using the following 5'
                   adaptor: GGCACGAG(G). Library constructed by Ling Hong
                   in the laboratory of Gerald M. Rubin (University of
                   California, Berkeley) using ZAP-cDNA synthesis kit
                   (Stratagene) and Superscript II RT (Life Technologies).
                   Note: this is a NIH MGC Library."
```

		24.9%; Score 672.6; DB 2; Length 934; Similarity 93.3%; Pred. No. 5.1e-133;	4
	Matches /4	7; Conservative 0; Mismatches 49; Indels 5; Gaps	4;
Q	•	GGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC	
Dł	2	GGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC	61
Qy	1131	TGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTC	1190
DŁ	62	${\tt TGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGGCTGGACTCAGATGTGGCTGACTC}$	121
Qς	1191	GTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCC	1250
Dł	122	GTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCC	181
Qy	1251	CCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCACCAGGGCAGTCT	1310
Dł	182	CCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCACCAGGGCAGTCT	241
Q5	1311	CTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAG	1370
Dł	242	CTGTCCCCGGCAGGATGGGCCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAG	301
Q۶	1371	CCCCTGGGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGA	1430
Dk	302	CCCCTGGGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGA	361
Q5	1431	GTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAA	1490
Dk	362	GTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAA	421
Q۶	1491	CATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTAT	1550
Dk	422	CATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAAT	481
Qy	7 1551	CAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCAC	1610
Dk	482	CAG-CTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCAC	540
Qy	7 1611	GCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTG	1670
Dł	541	GCTGCACAAGCCGGAAGACGTGAGGTTG-CCCTAGCTGGCTGTCAGACCCTGCTGAGTCC	599
Qy	1671	CATCGTTAGCTGTGGACCCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATG	1728
Dk	600	CATCGTTAGCTGTGGACCCCTGGCGTCCTGCTCAACCGGGCAGTCATCCCTGGCTAATG	659
Qy	1729	GACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGC	1788
Dł	660	GACCCACTGTGGGGAAGCCAGCCTGACAGTTGGAGCCTGGGCTCAAAAAGGAGTCGGTGC	719
Q	1789	GAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTAC	1848
Dk	720	GAGGGCAGCTGGGA-GATGTGCTGGACTGGGGGAAGAAGGGGGCCTCCACCCTCTATTAT	778

1849 TGCCAGCTGGAGGCCAGTGCC 1869 Qу 11 111 111111 11 11 779 TGGCAGTGGGAGGCAAGGCCC 799 Db

Search completed: March 6, 2005, 10:10:35 Job time : 8489.11 secs

GenCore version 5.1.6 Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

March 5, 2005, 15:11:26; Search time 11456.2 Seconds Run on:

(without alignments)

11407.261 Million cell updates/sec

US-10-624-932-1 COPY 46 2742 Title:

Perfect score: 2697

Sequence: 1 atggccgtccggcctgcct.....tgtcggaggctgagtgctga 2697

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

4708233 seqs, 24227607955 residues Searched:

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

GenEmbl:* Database :

> gb ba:* 1:

gb_htg:* 2:

3: gb in:*

4: gb om:*

gb ov:* 5:

6: gb pat:*

7: gb ph:*

gb pl:* 8:

9: gb pr:*

10: gb ro:*

11: gb sts:*

12: gb_sy:*

13: gb un:*

14: gb_vi:*

o.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2697	100.0	2752	6	AX449572	AX449572 Sequence
2	2687.4	99.6	2697	6	AX451652	AX451652 Sequence
3	2621.4	97.2	2881	6	AX527916	AX527916 Sequence

	0.405			_	~~~~	### ### ### ### ### ##################
4	2435.4	90.3	2784	6	CQ730306	CQ730306 Sequence
5	2343	86.9	3580	6	AX367094	AX367094 Sequence
6	2297	85.2	3992	10		AJ487852 Mus muscu
7	2252.2	83.5	2697	6	AX268596	AX268596 Sequence
8	2252.2	83.5	2697	10	RNU87305	U87305 Rattus norv
9	2252.2	83.5	3014	6	BD057524	BD057524 Netrin re
10	1957.4	72.6	3844	10		BC058084 Mus muscu
11	1625.4	60.3	2688	9	BC009333	BC009333 Homo sapi
12	1552.4	57.6	1787	6	BD057525	BD057525 Netrin re
13	1302.8	48.3	9700	6	AX054976	AX054976 Sequence
14	992	36.8	9299	10	MMU72634	U72634 Mus musculu
15	988.8	36.7	2962	5	AY187310	AY187310 Gallus ga
16	986	36.6	9328	10	AB118026	AB118026 Rattus no
17	970	36.0	3646	6	CQ881060	CQ881060 Sequence
18	970	36.0	3646	9	AF055634	AF055634 Homo sapi
19	942.4	34.9	3770	9	AY126437	AY126437 Homo sapi
20	936.2	34.7	2860	6	AX686445	AX686445 Sequence
21	936.2	34.7	2860	6	AX686447	AX686447 Sequence
22	912.8	33.8	3672	10	BC048162	BC048162 Mus muscu
23	912.8	33.8	3672	10	BC057560	BC057560 Mus muscu
24	912.2	33.8	2895	6	AX512281	AX512281 Sequence
25	904.2	33.5	2986	6	CQ881052	CQ881052 Sequence
26	904.2	33.5	3933	6	CQ850929	CQ850929 Sequence
27	904.2	33.5	3933	9	AK128132	AK128132 Homo sapi
28	904.2	33.5	3935	9	AB096256	AB096256 Homo sapi
29	901	33.4	3884	6	AR528525	AR528525 Sequence
30	901	33.4	3884	6	AX464012	AX464012 Sequence
31	901	33.4	3884	9	AY358351	AY358351 Homo sapi
32	892.4	33.1	2995	6	AX497288	AX497288 Sequence
33	892	33.1	4294	10	AK122575	AK122575 Mus muscu
34	874.6	32.4	3788	10	MMU487853	AJ487853 Mus muscu
35	852.4	31.6	2838	10		U87306 Rattus norv
36	843.6	31.3	2625	6	CQ721377	CQ721377 Sequence
37	841.4	31.2	2831	6	BD057526	BD057526 Netrin re
38	764.8	28.4	813	6	AX054892	AX054892 Sequence
39	751.4	27.9	4330	5	AY744919	AY744919 Petromyzo
40	738.4	27.4	2230	6	CQ845766	CQ845766 Sequence
41	738.4	27.4	2230	9	AK131380	AK131380 Homo sapi
42	717.2	26.6	2832	5	AY099459	AY099459 Xenopus 1
43	668.4	24.8	2612	6	CQ881064	CQ881064 Sequence
44	623.2	23.1	2661	6	AX800717	AX800717 Sequence
45	623.2	23.1	2868	6	AX800719	AX800719 Sequence
	020.2	20.1	2000	-	111000,15	12.000/15 bequeince

ALIGNMENTS

```
RESULT 1
AX449572
LOCUS
            AX449572
                                    2752 bp
                                               DNA
                                                       linear
                                                                PAT 03-JUL-2002
DEFINITION
           Sequence 1 from Patent W00210216.
            AX449572
ACCESSION
VERSION
            AX449572.1 GI:21698195
KEYWORDS
SOURCE
            Homo sapiens (human)
  ORGANISM Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
         Padigaru, M., Mezes, P., Mishra, V., Burgess, C., Casman, S.,
 AUTHORS
         Grosse, W.M., Alsobrook, J.P., Lepley, D.M., Gerlach, V.L.,
         Macdougall, J.R. and Smithson, G.
         Proteins and nucleic acids encoding same
 TITLE
         Patent: WO 0210216-A 1 07-FEB-2002;
 JOURNAL
         Curagen Corporation (US)
FEATURES
                Location/Qualifiers
                1. .2752
   source
                /organism="Homo sapiens"
                /mol type="unassigned DNA"
                /db xref="taxon:9606"
ORIGIN
 Query Match
                    100.0%;
                           Score 2697; DB 6;
                                          Length 2752;
 Best Local Similarity
                    100.0%; Pred. No. 0;
                         0; Mismatches
 Matches 2697; Conservative
                                                            0;
                                       0:
                                          Indels
                                                  0:
                                                     Gaps
         1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
Qу
          46 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTTGGCTC 105
Db
        61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
Qу
          106 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 165
Db
       121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
Qу
           166 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 225
Db
       181 GTGCTGCTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
Qу
          226 GTGCTGCTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 285
Db
       241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCACCAGACGGGAGCAGTGGGCTGCCC 300
Qу
           286 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCACCAGACGGGAGCAGTGGGCTGCCC 345
Db
       301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
Qу
          346 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 405
Db
       361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
Qу
           Db
       406 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCAACAAGAGTCAGAAGGCC 465
       421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
Qу
          466 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 525
Db
       481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
Qу
          Db
       526 TCCCTGGAGCAGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 585
Qу
       541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
```

	Db .	586	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	645
	Qy	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
	Db	646	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	705
	Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC	720
	Db ·	706	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTCATCGTCTAC	765
	Qу	721	GTGAACGGTGGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCC	780
	Db	766	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	825
	Qy	781	GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
	Db	826	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	885
	Qу	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC	900
	Db	886	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC	945
	Qу	901	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	960
	Db	946	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	1005
	Qу	961	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGGAGTGCCAGGGCACTGACCTGGAC	1020
	Db	1006	GAGTGCTCTGACCCAGCACCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGAC	1065
	Qy	1021	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC	1080
	Db	1066	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC	1125
	Qy	1081	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGCTGCTGCTGCTGCTCATC	1140
	Db	1126	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGCTGCTGCTGCTCATC	1185
	Qy	1141	CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
	Db	1186	CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1245
	Qy	1201	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
•	Db	1246	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1305
	Qy	1261	ACCATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCCCGG	1320
	Db	1306	ACCATCCAGCCGGACCTCAGCACCACCACCACCAGGGCAGTCTCTGTCCCCGG	1365
	Qy	1321	CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1380
	Db	1366	CAGGATGGGCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1425
	Qy	1381	GGCGGCCGCACACTGCACCACACTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCC	1440
	Db	1426	GGCGGCCGCACACACACACACACACCTCTGAGGCCGAGGAGTTCGTCTCC	1485

	Qy	1441	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1500
·	Db	1486	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1545
	Qy	1501	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1560
	Db	1546	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1605
	Qy	1561	ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
	Db	1606	ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1665
	Qy	1621	CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC	1680
	Db	1666	CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC	1725
	Qу	1681	TGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1740
	Db	1726	TGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1785
	Qу	1741	GAGCCCAGCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
	Db	1786	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1845
	Qy .	1801	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG	1860
	Db	1846	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG	1905
	Qу	1861	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	1920
	Db	1906	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	1965
	Qу	1921	CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	1980
	Db	1966	CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	2025
	Qу	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2040
	Ďb	2026	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2085
	Qу	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC	2100
	Db	2086	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC	2145
	Qу	2101	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG	2160
	Db	2146	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG	2205
	Qу	2161	AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG	2220
	Db	2206	AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG	2265
	Qу	2221	CAGCGGTACTTGCACTGCACCTTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG	2280
	Db	2266	CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG	2325

```
2281 GCCTGCAAGCTGTGGGTGGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340
Qy
          2326 GCCTGCAAGCTGTGGGTGGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2385
Db
      2341 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400
Qy
          2386 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2445
Db
      2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460
Qу
          2446 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2505
Db
      2461 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
Qy
          2506 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2565
Db
      2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTC 2580
Qy
          2566 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2625
Db
      Qу
          Db
      2641 GCTGGACTGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
Qу
          2686 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2742
Db
RESULT 2
AX451652
         AX451652
                           2697 bp
                                    DNA
                                          linear
                                                 PAT 03-JUL-2002
LOCUS
         Sequence 1 from Patent WO0233080.
DEFINITION
ACCESSION
         AX451652
         AX451652.1 GI:21698587
VERSION
KEYWORDS
SOURCE
         Homo sapiens (human)
 ORGANISM
         Homo sapiens
         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
         Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
 AUTHORS
         Koehler, R.H.
         Regulation of human netrin binding membrane receptor unc5h-1
 TITLE
         Patent: WO 0233080-A 1 25-APR-2002;
 JOURNAL
         Bayer Aktiengesellschaft (DE)
FEATURES
                Location/Qualifiers
                1. .2697
    source
                /organism="Homo sapiens"
                /mol type="unassigned DNA"
                /db xref="taxon:9606"
ORIGIN
                    99.68;
                          Score 2687.4;
                                     DB 6;
                                          Length 2697;
 Query Match
                    99.8%;
 Best Local Similarity
                          Pred. No. 0;
                           Mismatches
                                                    Gaps
                                                           0;
 Matches 2691; Conservative
                         0;
                                       6;
                                          Indels
                                                 0;
```

Qу

Db	1		60
Qy	61	CGCGGCTCGGGTGCCCAGCAGAGTGCCACCCGTGGCCAACCCAGTGCCTGGTGCCAACCCG	120
Db	61		120
Qу	121	${\tt GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA}$	180
Db	121		180
Qу	181	GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG	240
Db	181	GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG	240
Qу	241	TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC	300
Db	241	TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC	300
Qу	301	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG	360
Db	301	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG	360
Qу	361	GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	420
Db	361	GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	420
Qу	421	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
Db	421	TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
Qу	481	TCCCTGGAGCAGGCATCGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Db	481	TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Qу	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	600
Db	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATC	600
Qy (601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTÄCACC	660
Db	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
Qу	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC	720
Db	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTCATCGTCTAC	720
Qу	721	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	780
Db	721	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCC	780
Qy	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
Db	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
Qу	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC	900

Db	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTGGACGGCAGC	900
Qy	901	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	960
Db	901	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	960
Qy	961	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGGG	1020
Db	961	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGAC	1020
Qу	1021	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC	1080
Db	1021	ACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTGGCC	1080
Qу	1081	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATC	1140
Db	1081	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGCTGCTGCTGCTCATC	1140
Qy	1141	CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
Db	1141	CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
Qy	1201	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
Db	1201	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
Qy	1261	ACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG	1320
Db	1261	ACCATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCCCGG	1320
Qy	1321	CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1380
Db	1321	CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1380
Qу	1381	GGCGGCCGCCACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC	1440
Db	1381	GGCGGCCGCCACACTGCACCACACTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC	1440
Qу	1441	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1500
Db	1441	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1500
Qy	1501	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1560
Db	1501	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTC	1560
Qy	1561	ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Db	1561	${\tt ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG}$	1620
Qy	1621	CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC	1680
Db	1621	CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC	1680
Qy		TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	
Db	1681	TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1740

Qy	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Db ·	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Qy	1801	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG	1860
Db	1801	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG	1860
Qу	1861	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	1920
Db	1861	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	1920
Qу	1921	CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	1980
Db	1921	CTCAGCGTGGCTGCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	1980
Qу	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2040
Db	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2040
Qy	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC	2100
Db	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC	2100
Qy	2101	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG	2160
Db	2101	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG	2160
Qy	2161	AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG	2220
Db	2161	AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG	2220
Qy	2221	CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG	2280
Db	2221	CAGCGGTACTTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG	2280
Qy	2281	GCCTGCAAGCTGTGGGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2340
Db	2281	GCCTGCAAGCTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2340
Qy	2341	AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC	2400
Db	2341	AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC	2400
Qy	2401	CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT	2460
Db	2401	CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT	2460
Qу	2461	TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC	2520
Db	2461	TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC	2520
Qу	2521	CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTC	2.58.0
Db	2521	CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTC	2580

```
Qy
          Db
      2641 GCTGGACTGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
Qу
          2641 GCTGGACTGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
Db
RESULT 3
AX527916
                                                 PAT 21-NOV-2002
LOCUS
         AX527916
                           2881 bp
                                    DNA
                                          linear
         Sequence 1 from Patent W00229038.
DEFINITION
ACCESSION
         AX527916
         AX527916.1 GI:25172359
VERSION
KEYWORDS
SOURCE
         Homo sapiens (human)
 ORGANISM
        Homo sapiens
         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
         Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
 AUTHORS
         Herrmann, J.L., Rastelli, L. and Shimkets, R.A.
         Novel proteins and nucleic acids encoding same and antibodies
 TITLE
         directed against these proteins
         Patent: WO 0229038-A 1 11-APR-2002;
 JOURNAL
         Curagen Corporation (US)
FEATURES
                Location/Qualifiers
                1. .2881
   source
                /organism="Homo sapiens"
                /mol type="unassigned DNA"
                /db xref="taxon:9606"
ORIGIN
                    97.2%;
                          Score 2621.4;
                                          Length 2881;
 Query Match
                                     DB 6;
 Best Local Similarity
                   98.9%;
                          Pred. No. 0;
 Matches 2673: Conservative
                         0; Mismatches
                                      21;
                                          Indels
                                                    Gaps
                                                           3:
         1 ATGGCCGTCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
Qy
          87 ATGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGCCATAGTCCTCGCCGCTTGGCTC 146
Db
        61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
Qу
          147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206
Db
       121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
Qy
          207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266
Dh
       181 GTGCTGCTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
Qу
          267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326
Db
       241 TGGGTGCGCCAGGTGACCACGTGATCGAGCGCAGCAGACGGGAGCAGTGGGCTGCCC 300
Qу
          327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386
Db
```

Qy	301	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG	360
Db	387	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG	446
Qу	361	GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	420
Db	447	GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	506
Qу	421	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
Db	507	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	566
Qy .	481	TCCCTGGAGCAGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Db	567	TCCCTGGAGCAGGCATCGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	626
Qу	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	600
Db	627	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATC	686
Qу	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
Db.	687	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	746
Qу	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC	720
Db	747	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC	806
Qу	721	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	780
Db	807	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	866
Qу	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
Db	867	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	926
Qу	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC	897
Db	927	TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC	986
Qу	898	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	957
Db	987	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	1046
Qу	958	CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG	1017
Db	1047	CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG	1106
Qy	1018	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1077
Db	1107	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1166
Qу	1078	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC	1137
Db	1167	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC	1226
017	1139	λ TCCTCCTTTT λ TTCCCCCC λ AC λ AC λ ACC λ CCCCCTCC λ CTC λ CTCC λ CTCC λ CTC λ CTCC λ CTC	1107

•

-

Db	1227	ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1286
Qу	1198	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1257
Db	1287	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1346
Qу	1258	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCC	1317
Db ·	1347	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCC	1403
Qу	1318	CGGCAGGATGGGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1377
Db	1404	CGGCAGGATGGGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1463
Qу	1378	GGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1437
Db	1464	GGTGGCGGCCACACACTGCACCACACCTCTGAGGCCGAGGAGTTCGTC	1523
Qу	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1497
Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1583
Qу	1498	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1557
Db	1584	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1643
Qу	1558	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1617
Db	1644	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1703
Qу	1618	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1677
Db	1704	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1763
Qy .	1678	AGCTGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1737
Db	1764	AGCTGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1823
Qу	1738	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1797
Db	1824	GGGGAGCCCAGCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1883
Qу	1798	TGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1854
Db	1884	TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1943
Qу	1855	CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	1914
Db	1944	CTGGAGGCCAGTGCCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	2003
Qу	1915	GAGGCCCTCAGCGTGGCTGCCCCCAAGCCCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	1974
Db	2004	GAGGCCCTCAGCGTGGCTGCCCCAAGCCCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	2063
Qу	1975	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2034

Db	2064	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2123
Qу	2035	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC	2094
Db	2124	AAGGAGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC	2183
Qy	2095	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	2154
Db	2184	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	2243
Qy	2155	CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT	2214
Db	2244	CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT	2303
Qу	2215	GGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT	2274
Db .	2304	GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT	2363
Qy	2275	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC	2334
Db	2364	GACCTGGCCTGCAAGCTGTGGGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC	2423
Qy	2335	AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2394
Db	2424	AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2483
QУ	2395	GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG	2454
Db	2484	GGGGTCCCAGCCCTGGTGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG	2543
Qy	2455	ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG	2514
Db	2544	ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG	2603
Qу	2515	AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATG	2574
Dp ´	2604	AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATG	2663
QУ	2575	ATCCTCAACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC	2634
Db	2664	ATCCTCAACCTGTGGGAGGCGCGCCACTTCCCCCAACGGCAACCTCAGCCAGC	2723
Qу	2635	GCAGTGGCTGGACTGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC	2694
Db	2724	GCAGTGGCTGGACTGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC	2783
QУ	2695	TGA 2697	
Db	2784	TGA 2786	

RESULT 4 CQ730306

LOCUS CQ730306 2784 bp DNA linear PAT 03-FEB-2004

DEFINITION Sequence 16240 from Patent WO02068579.

ACCESSION CQ730306

VERSION CQ730306.1 GI:42303801

```
KEYWORDS
SOURCE
         Homo sapiens (human)
 ORGANISM
         Homo sapiens
         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
         Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
 AUTHORS
         Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
         Kits, such as nucleic acid arrays, comprising a majority of
 TITLE
         humanexons or transcripts, for detecting expression and other uses
         thereof
         Patent: WO 02068579-A 16240 06-SEP-2002;
 JOURNAL
         PE Corporation (NY) (US)
FEATURES
                 Location/Qualifiers
                 1. .2784
    source
                 /organism="Homo sapiens"
                 /mol type="unassigned DNA"
                 /db xref="taxon:9606"
ORIGIN
 Query Match
                     90.3%;
                           Score 2435.4;
                                        DB 6;
                                             Length 2784;
 Best Local Similarity
                     93.5%;
                           Pred. No. 0;
 Matches 2646; Conservative
                           0; Mismatches
                                                               3;
                                         6;
                                             Indels 177;
                                                        Gaps
          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
Qу
           1 ATGGCCGTCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
Db
         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
Qу
           61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
Db
        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
Qу
           121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
Db
        181 GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
Qу
           181 GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
Db
        241 TGGGTGCGCCAGGTGACCACGTGATCGAGCGCAGCAGACGGGAGCAGTGGGCTGCCC 300
Qу
           241 TGGGTGCGCCAGGTGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
Db
        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
Qу
           301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
Db
        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
Qу
           361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
Db
        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
Qy
           11111111111
                          421 TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
Db
        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
Qy
```

	Db	481	TCCCTGGAGCAGGCATCGTGCCGTGCCGTCCACCGGAGGCCATCCCTCCAGCCGAG	540
	Qy	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	600
	Db	541		600
	Qy	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
	Db	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
	Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC	720
	Db	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTCATCGTCTAC	720
	Qy	721		720
	Db	721	${\tt GGTGGGCCCCGGGACTCCCTGGTCACAGGGAGAGGCACTGCGGTGCCCCTGGGCAGTGAC}$	780
	Qу	721	GTGAACGGTGGGTCGACGTGGACCGAGTGG	753
	Db	781	ATGTGGCTGTCCTGTCCGGCCAGTGAACGGTGGGTCGACGTGGACCGAGTGG	840
	Qу	754	TCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAAC	813
·	Db	841	TCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAAC	900
	Qу	814	CCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGC	873
	Db	901	CCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGC	960
	Qу	874	GCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGG	933
	Db	961	GCCACCCTGTGCCCAGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGG	1020
	Qy	934	CTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGG	993
	Db	1021	CTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGG	1080
	QУ	994	GAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACAC	1053
	Db	1081	GAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACAC	1140
	Qу	1054	AGTGCTTCTGGCCCT	1068
	Db	1141	AACTCCTACACCCCTGCCCCCACCAAGGCCATGCTGTCTCCCGCAGCTGCTTCTGGCCCT	1200
	Qy	1069	GAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC	1128
	Db	1201	GAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC	1260
	Qy	1129	CTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGAC	1188
	Db	1261	CTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGAC	1320
	Qу	1189	TCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAAC	1248
	Db	1321	TCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAAC	1380

.

Qу	1249	CCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCACCACCAGGGCAGT	1308
Db	1381	CCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCACCAGGGCAGT	1440
Qy	1309	CTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTC	1368
Db	1441	CTCTGTCCCCGGCAGGATGGGCCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTC	1500
Qу	1369	AGCCCCTGGGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAG	1428
Db	1501	AGCCCCTGGGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAG	1560
Qу	1429	GAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGC	1488
Db	1561	GAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGC	1620
Qу	1489	AACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGT	1548
Db	1621	AACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGA	1680
Qу	1549	ATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTC	1608
Db	1681	ATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTC	1740
Qу	1609	ACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTG	1668
Db	1741	ACGCTGCACAAGCCGGAAGACGTG	1764
Qу	1669	CCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATG	1728
Db	1765	AGCTGTGGACCCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATG	1815
Qу	1729	GACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGC	1788
Db	1816	GACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGC	1875
Qу	1789	GAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTAC	1848
Db	1876	GAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTAC	1935
Qу	1849	TGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTG	1908
Db	1936	TGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTG	1995
Qy ·	1909	GTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCG	1968
Db	1996	GTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCG	2055
Qу	1969	GTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2028
Db	2056	GTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2115
Qу	2029	GCACTCAAGGAGGTGCTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCA	2088
Db	2116		2175

Qy Db		CGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC
Qy	2149	AGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATC 2208
Db	223	
Qу	2209	TGGAATGGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGC 2268
Db	229	
Qy	226	ACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTC 2328
Db	235	
Qy	2329	AGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGT 2388
Db	241	AGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGT 2475
Qу	2389	GAAGCGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGG 2448
Db	2476	GAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGG 2535
Qу	2449	CAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTG 2508
Db	2536	5 CAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTG 2595
Qу	2509	GCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACA 2568
Db	259	
Qy	2569	GCCATGATCCTCAACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC
Db	265	
Qу	2629	GCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCT 2688
Db	271	
Qy	2689	GAGTGCTGA 2697
Db	277	
RESULT 5 AX367094 LOCUS DEFINITION ACCESSION KEYWORDS	ON S	AX367094 3580 bp DNA linear PAT 16-FEB-2002 Sequence 13 from Patent WO0198354. AX367094 AX367094.1 GI:18855296
SOURCE ORGANI REFERENC	I SM I I	Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```
Griffin, J.A., Kallick, D.A., Tribouley, C.M., Yue, H., Nguyen, D.B.,
 AUTHORS
         Tang, Y.T., Lal, P., Policky, J.L., Azimzai, Y., Lu, D.A., Graul, R.,
         Yao, M.G., Burford, N., Hafalia, A.J., Baughn, M.R., Bandman, O.,
         Patterson, C., Yang, J., Xu, Y., Warren, B.A., Ding, L. and
         Sanjanwala, M.S.
         Receptors
 TITLE
 JOURNAL
         Patent: WO 0198354-A 13 27-DEC-2001;
         Incyte Genomics, Inc. (US)
FEATURES
                Location/Qualifiers
                 1. .3580
    source
                 /organism="Homo sapiens"
                 /mol type="unassigned DNA"
                 /db xref="taxon:9606"
                 /note="Incyte ID No: 6052371CB1"
ORIGIN
                     86.9%;
                           Score 2343; DB 6; Length 3580;
 Query Match
                     93.6%;
                           Pred. No. 0;
 Best Local Similarity
 Matches 2524; Conservative
                          0: Mismatches
                                         5;
                                            Indels 168:
         1 ATGGCCGTCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
Qy
           4 ATGGCCGTCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 63
Db
        61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
Qу
           64 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 123
Db
        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
Qу
           124 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 183
Db
        181 GTGCTGCTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
Qу
           184 GTGCTGCTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 243
Db
        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
Qу
           244 TGGGTGCGCCAGGTGACCACGTGATCGAGCGCAGCAGACGGGAGCAGTGGGCTGCCC 303
Db
        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
Qу
           304 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 363
Db
        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
Qу
           364 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 423
Db
        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
Qу
           1111111111111
                         424 TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 483
Db
        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
Qy
           484 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 543
Db
        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
Qy
```

Db	544	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	603
Qу	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
Db	604	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	663
Qу	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC	720
Db	664	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTAC	723
Qу	721	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCC	780
Db	724		726
Qу	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
Db	727		726
Qу	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC	900
Db	. 727	GACGGCAGC	735
Qу	901	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	960
Db	736	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	795
Qу	961	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGGAGTGCCAGGGCACTGACCTGGAC	1020
Db	796	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGAC	855
Qу	1021	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC	1080
Db	856	ACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTGGCC	915
Qу	1081	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATC	1140
Db	916	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGCTGCTGCTGCTCATC	975
Qу	1141	CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	12'00
Db	976	CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1035
Qу	1201	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
Db	1036	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1095
Qу	1261	ACCATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCCCGG	1320
Db	1096	ACCATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCCCGG	1155
Qу	1321	CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1380
Db	1156	CAGGATGGGCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1215
Qу	1381	GGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC	1440

Db .	1216	${\tt GGCGGCCGCCACACACTGCACCACCTCTGAGGCCGAGGAGTTCGTCTCC}$	1275
Qу		CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	
Db	1276	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1335
Qу	1501	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1560
Db	1336	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTC	1395
Qy	1561	${\tt ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG}$	1620
Db	1396		1455
Qy	1621	$\tt CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC$	1680
Db	1456		1515
Qу	1681	TGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1740
Db	1516	TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1575
Qy .	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Db	1576		1635
Qy	1801	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG	1860
Db	1636		1695
Qy	1861	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	1920
Db	1696	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	1755
Qy	1921	CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	1980
Db	1756	CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	1815
Qy	1981	${\tt TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG}$	2040
Db	1816		1875
Qy	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC	2100
Db	1876		1935
Qу	2101	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG	2160
Db	1936	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG	1995
Qy	2161	AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG	2220
Db	1996		2055
Qу	2221	CAGCGGTACTTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG	2280
Db	2056		2115

.

Qy	228	1 GCCTGCAAGCTGTGGGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340
Db	211	6 GCCTGCAAGCTGTGGGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2175
Qу	234	1 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400
Db	217	
Qy .	240	1 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460
Db	223	6 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2295
Qу	246	1 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
Db	229	6 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2355
Qy	252	1 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580
Db ·	235	6 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTC 2415
Qy	258	AACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC
Db	241	6 AACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC
Qу	264	1 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
Db	247	6 GCTGGACTGGGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2532
RESULT 6 MMU487852 LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANIS REFERENCE AUTHORS TITLE JOURNAL MEDLINE PUBMEN REFERENCE AUTHORS TITLE JOURNAL FEATURES	ON N SM E S L E D E S	MMU487852 Mus musculus mRNA for netrin receptor Unc5h1 (Unc5h1 gene). AJ487852 AJ487852.1 GI:22035783 Mus musculus (house mouse) Mus musculus (house mouse) Mus musculus (house mouse) Mus musculus (Eutheria; Rodentia; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Engelkamp, D. Cloning of three mouse Unc5 genes and their expression patterns at mid-gestation Mech. Dev. 118 (1-2), 191-197 (2002) 22239710 12351186 2 (bases 1 to 3992) Engelkamp, D. Direct Submission Submitted (15-MAY-2002) Neuroanatomy, Max Planck Institute for Brain Research, Deutschordenstrasse 46, Frankfurt 60528, GERMANY Location/Qualifiers
sou	rce	<pre>13992 /organism="Mus musculus" /mol_type="mRNA"</pre>

translation="MAVRPGLWPALLGIVLTAWLRGSGAQQSATVANPVPGANPDLLP/ HFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTM EVRINVSRQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEV SLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTAN YTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLN GGAFCEGONVOKTACATLCPVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEEC RGADLDTRNCTSDLCLHTSSGPEDVALYIGLVAVAVCLILLLLVLVLIYCRKKEGLDS DVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQDGPSPKFQL SNGHLLSPLGSGRHTLHHSSPTSEAEDFVSRLSTONYFRSLPRGTSNMAYGTFNFLGG RLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGV LLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEDVLHLGEESPSHLYYCQLEAGAC YVFTEQLGRFALVGEALSVAATKRLRLLLFAPVACTSLEYNIRVYCLHDTHDALKEVV QLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGT QQYLHCTFTLERVNASTSDLACKVWVWQVEGDGQSFNINFNITKDTRFAEMLALESEG GVPALVGPSAFKIPFLIRQKIITSLDPPCSRGADWRTLAQKLHLDSHLSFFASKPSPT AMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC"

ORIGIN

Query Ma Best Loc	cal :	Simi		ty	9		} ;	Pr	ed.	No	э.	0;		10	•		ngt		399			_	•
Matches	244	7;	Cons	erva	ati	ve		0;	Mi	sma	atc	hes	5	250	;	In	del	s		0;	C	Saps	0;
Qу	1		GCCG																GC(CGCI		GCTC	60
Db	232																		AC'	rĠĊŒ	٠.	GCTT	291
Qy	61		GGCT																			ACCCG	120
Db	292									, .											٠.	ACCCA	351
Qy	121		CTGC			ACTT																AGCCA	180
Db	352	GAC	CTGC																			AGCCC	411
Qy	181		CTGC																			GGAG	240
Db	412	GTG	CTGC	TGG:	rgt	GCA/	AGG	CTG	TGC	CC	GCC	ACC	CA	GAT	'CT'	CT	TCA	AG'	rgo	CAAC	GG	GGAA	471
Qу	241		GTGC					ACG		TC				CAC		ACG		GC.		rggo II		GCCC	300
Db	472	TGG	GTTC	GCC	AGG'	TCG	ATC.	ACG	TCA	TC	GAA	CGC	AG	CAC	TGA	ACG	GCA	GC.	AGO	CGGZ	\TT	GCCA	531
Qу	301		ATGG				TTA	ATG 			ÄGG							TĢʻ	TTC	CGGC	C7	GGAG	360
Db	532	ACC	ATGG	AGG'	rcc	GGA'	ГСА	ACG	TAT	'CA	AGG	CAC	CA	GGT	'CG/	AGA	AAG	TG'	rT7	rggc	C	GGAG	591

	Qу	361	GAATACTGGTGCCAGTGGGTGGGATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	420
	Db	592		651
	QУ	421	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
	Db	652	TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCGCTGGCCAAGGAAGTG	711
	Qу	481	TCCCTGGAGCAGGGCATCGTGCCCTGCCGTCCACCGGAGGCCATCCCTCCAGCCGAG	540
	Db	712		771
	Qу	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATC	600
	Db	772		831
	Qy	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
	Db	832		891
	Qу	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC	720
-	Db	892	TGCGTGGCCAAGAACATCGTGGCCCGTCGCCGAAGCGCCTCTGCGGCCGTCATTGTTTAT	951
	Qy	721	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCC	780
	Db	952		1011
	Qу	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
	Db	1012	GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC	1071
	Qу	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC	900
	Db	1072	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGGAGC	1131
	Qу	901	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	960
	Db	1132	TGGAGCCCATGGAGTAAGTGGTCAGCCTGCGGGCTTGACTGCACCCACTGGCGGAGCCGG	1191
	Qу	961	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC	1020
	Db	1192	GAGTGCTCCGACCCAGCGCCCCGCAACGGAGGTGAGGAGTGCCGGGGGTGCTGACCTGGAC	1251
	Qу	1021	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC	1080
	Db	1252	ACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCTCTTCCGGCCCCGAGGACGTGGCT	1311
	Qy	1081	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATC	1140
	Db	1312	CTCTACATCGGCCTCGTCGCCGTGGCCGTGTCCTCATCTTGCTGCTGCTGGTCCTCGTC	1371
	Qy	1141	CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
	Db	1372	CTCATCTACTGCCGCAAGAAGGAAGGACTGGACTCAGACGTGGCTGACTCATCCATC	1431
	Ov	1201	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260

Db	1432	ACCTCAGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAATCCCCATCTGCTC	1491
Qу	1261	ACCATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCCCGG	1320
Db	1492	ACCATCCAACCGGACCTCAGCACCACCACCACCTACCAGGGCAGCCTGTGTCCCCGG	1551
Qу	1321	CÁGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1380
Db	1552	CAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTGGGC	1611
Qу	1381	GGCGGCCACACACTGCACCACACTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCC	1440
Db	1612	AGTGGCCGCCATACGCTGCACCACAGCTCCCCCACCTCTGAGGCTGAGGACTTCGTCTCC	1671
Qу	1441	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1500
Db	1672	CGCCTCTCCACCCAAAACTACTTTCGTTCTCTGCCCCGCGGTACCAGCAACATGGCCTAT	1731
QУ	1501	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1560
Db	1732	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAACACAGGAATCAGCCTCCTC	1791
Qу	1561	ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Db	1792	ATACCCCGGACGCCATCCCCGAGGAAAGATCTACGAGATCTACCTCACTCTGCACAAG	1851
Qу	1621	CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC	1680
Db	1852	CCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCTATCGTTAGC	1911
Qу	1681	TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1740
Db	1912	TGTGGGCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCCATGGACCACTGCGGG	1971
Qу	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Db	1972	GAGCCCAGTCCCGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCCTGTGAGGGCAGCTGG	2031
Qу	1801	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG	1860
Db	2032	GAGGACGTGCTGCACCTTGGTGAGGAGTCGCCCTCTCATCTCTACTACTGCCAGCTGGAG	2091
Qу	1861	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	1920
Db .	2092	GCCGGGGCCTGCTATGTCTTCACCGAGCAGCTAGGCCGCTTTGCCCTGGTGGGAGAGGCC	2151
Qy	1921	CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	1980
Db	2152	CTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCCCCTGTGGCCTGTACG	2211
Qу	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2040
Db .	2212	TCCCTCGAGTACAACATCCGAGTGTACTGCCTGCACGACACCCACGATGCTCTCAAGGAG	2271
Qу	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC	2100

```
Db
      2272 GTGGTGCAGCTGGAGAAGCAGCTGGGTGGACAGCTGATCCAGGAGCCCCGTGTCCTGCAC 2331
      2101 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2160
Qv
          2332 TTCAAAGACAGTTACCACAACCTACGTCTGTCCATCCACGACGTGCCCAGCTCCCTGTGG 2391
Db
      2161 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2220
Qy
          2392 AAGAGCAAGCTCCTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAATGGCACT 2451
Db
      2221 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2280
Qу
                 111111 11 11111
      2452 CAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCGTCAATGCCAGCACCAGCGACCTG 2511
Db
      2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340
Qу
          2512 GCCTGCAAGGTGTGGGTGGGCAGGTGGAGGGAGACAGAGCTTCAACATCAACTTT 2571
Db
      2341 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400
Qy
          2572 AACATCACTAAGGACACGAGGTTTGCTGAAATGCTGGCTCTGGAGAGTGAAGGGGGGGTC 2631
Db
      2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460
Qy
          2632 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAGATCATT 2691
Db
      2461 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
Qy
           2692 ACCAGCCTGGACCCACCCTGCAGCCGGGGGCGCCGACTGGCGAACTCTAGCCCAGAAACTT 2751
Db
      2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCACAGCCACAGCCATGATCCTC 2580
Qy
          2752 CACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATGATCCTC 2811
Db
      Qy
          2812 AACCTATGGGAGGCGCGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCCGCAGCTGTG 2871
Db
      2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
Qy
          2872 GCCGGACTGGGCCAGATGCTGGCCTCTTCACCGTGTCAGAGGCCGAGTGCTGA 2928
Db
RESULT 7
AX268596
LOCUS
                           2697 bp
                                                PAT 29-OCT-2001
         AX268596
                                   DNA
                                         linear
         Sequence 15 from Patent WO0175440.
DEFINITION
ACCESSION
         AX268596
         AX268596.1 GI:16541710
VERSION
KEYWORDS
SOURCE
         Rattus sp.
 ORGANISM
         Rattus sp.
         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
         Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
         Rattus.
```

Cochran, S.W., Paterson, G.Y., Ohashi, Y.W., Morris, B.Y. and

REFERENCE

AUTHORS

Pratt, J.Y.

TITLE Schizophrenia related genes

JOURNAL Patent: WO 0175440-A 15 11-OCT-2001;

WELFIDE CORPORATION (JP)

FEATURES Location/Qualifiers

source 1. .2697

/organism="Rattus sp."
/mol_type="unassigned DNA"
/db_xref="taxon:10118"

ORIGIN

	cal	83.5%; Similarity 89.7%; P G; Conservative 0;	red.	No. 0;	DB 6;	Length Indels	2697; 0;	Gaps	0;
QУ	1	ATGGCCGTCCGGCCCGGCCTG							60
Db	1	ATGGCCGTCCGGCCCGGCCTG							60
Qy.	61	CGCGGCTCGGGTGCCCAGCAG.							120
Db	61	CGTGGTTCGGGTGCCCAGCAG							120
Qу	121	GACCTGCTTCCCCACTTCCTG							180
Db	121	GACCTGCTGCCCCACTTCCTG							180
Qу	181	GTGCTGCTTGTGTGCAAGGCC							240
Db	181	GTGTTGTTGGTGCAAGGCT							240
Qу	241	TGGGTGCGCCAGGTGGACCAC						CTGCCC	300
Db ·	241	TGGGTCCGCCAGGTCGATCAC							300
Qу	301	ACCATGGAGGTCCGCATTAAT							360
Db	301	ACCATGGAGGTCCGTATCAAC							360
Qу	361	GAATACTGGTGCCAGTGCGTG							420
Db	361	GAATACTGGTGCCAGTGTGTG							420
Qy	421	TACATCCGCATAGCCAGATTG							480
Db	421	TACATCCGGATTGCCTATTTG							480
Qy	481	TCCCTGGAGCAGGGCATCGTG							540
Db	481								540
Qy	541	GTGGAGTGGCTCCGGAACGAG							600
Db	541								600
Qy	601	ACGCGGGAGCACAGCCTGGTG							660

Db	601	${\tt ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC}$	660
Qу	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTAC	720
Db	661	TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT	720
Qу	721	GTGAACGGTGGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	780
Db	721	GTGAACGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT	780
Qy .	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
Db	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCCTTC	840
Qy	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC	900
Db	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGGAGC	900
Qу	901	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	960
Db	901	TGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGCCGC	960
Qy	961	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC	1020
Db	961	GAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTGGAC	1020
Qy	1021	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC	1080
Db	1021	ACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTGGCT	1080
Qу	1081	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGCTGCTGCTGCTGCTCATC	1140
Db	1081	CTCTACATCGGCCTTGTCGCTGTGCCTCTTCTTGCTGTTGCTGGCCCTTTGA	1140
Qу	1141	CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
Db	1141	CTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATCCTC	1200
Qу	1201	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACACCCCCATCTGCTC	1260
Db	1201	ACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTGCTC	1260
Qу	1261	ACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG	1320
Db	1261	ACCATCCAGCCAGACCTCAGCACCACCACCTACCAGGGCAGTCTATGTTCGAGG	1320
Qу	1321	CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1380
Db	1321	CAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTGGGG	1380
Qу	1381	GGCGGCCACACACTGCACCACACTCTCACCTCTGAGGCCGAGGAGTTCGTCTCC	1440
Db	1381	AGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTCTCC	1440
Qy	1441	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1500
Db	1441	CGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCGCGGCACCAGCAACATGGCCTAC	1500

	Qу	1501 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC 1560
	Db	1501 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTCCTC 1560
	Qу	1561 ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1620
,	Db	1561 ATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCACAAG 1620
	Qу	1621 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1680
	Db	
	Qу	1681 TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1740
	Db	
	Qу	1741 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1800
	Db	
	Qу	1801 GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1860
	Db	1801 GAGGATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAGCTGGAG 1860
	Qу	1861 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 1920
	Db	1861 GCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGAGAGGCC 1920
	Qу	1921 CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 1980
	Db	1921 CTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCCTGTACG 1980
	Qу	1981 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2040
•	Db	1981 TCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCCACGACGCTCTCAAGGAG 2040
	Qу	2041 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2100
	Db	2041 GTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTCCTGCAC 2100
	Qу	2101 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2160
	Db	2101 TTCAAAGACAGTTACCACAACCTACGTCTCCATCCACGACGTGCCCAGCTCCCTGTGG 2160
	Qу	2161 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2220
	Db	2161 AAGAGCAAGCTACTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAACGGCACC 2220
	Qу	2221 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2280
	Db	2221 CAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGCGACCTG 2280
	Qу	2281 GCCTGCAAGCTGTGGGTGGGGGGGGGGGGGGGGGGGGGG
	Db	2281 GCCTGCAAGGTGTGGCTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACTTC 2340

CDS

/dev_stage="18 day embyo and 13 day embryo"
1. .2697

/codon start=1

/product="transmembrane receptor UNC5H1"

/protein_id="AAB57678.1"

/db xref="GI:2055392"

/translation="MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLP ${\tt HFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTM}$ EVRINVSROOVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEV SLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTAN YTCVAKNIVARRRSTSAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLN GGAFCEGONVOKTACATLCPVDGSWSSWSKWSACGLDCTHWRSRECSDPAPRNGGEEC RGADLDTRNCTSDLCLHTASCPEDVALYIGLVAVAVCLFLLLLALGLIYCRKKEGLDS ${\tt DVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCSRQDGPSPKFQL}$ SNGHLLSPLGSGRHTLHHSSPTSEAEDFVSRLSTQNYFRSLPRGTSNMAYGTFNFLGG RLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPVVSCGPPGV LLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEDVLHLGEESPSHLYYCQLEAGAC YVFTEQLGRFALVGEALSVAATKRLRLLLFAPVACTSLEYNIRVYCLHDTHDALKEVV QLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGT QQYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNINFNITKDTRFAELLALESEG GVPALVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQKLHLDSHLSFFASKPSPT AMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC"

83.5%; Score 2252.2; DB 10; Length 2697;

ORIGIN

Query Match

Quely I			0.30, D	and M	_ 0.	22 10,	Dengen.	203.	•	
		Similarity 8 9; Conservati				278;	Indels	0;	Gaps	0;
Qy	1	ATGGCCGTCCGGC								60
Db	1	ATGGCCGTCCGGC								60
Qу	61	CGCGGCTCGGGTG								120
Db	61	CGTGGTTCGGGTG								120
Qу	121	GACCTGCTTCCCC								180
Db	121	GACCTGCTGCCCC								180
Qу	181	GTGCTGCTTGTGT								240
Db	181	GTGTTGTTGGTGT								240
Qу	241	TGGGTGCGCCAGG							CTGCCC	300
Db	241	TGGGTCCGCCAGG								300
Qу	301	ACCATGGAGGTCC								360
Db	301	ACCATGGAGGTCC								360
Qy	361	GAATACTGGTGCC								420
Db	361	GAATACTGGTGCC								420
Qy	421	TACATCCGCATAC	CCAGATTG	CGCAAG.	AACTTCG.	AGCAGG/	AGCCGCTGG	CCAAG	GAGGTG	480

Db	421	TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG	480
Qy	481	TCCCTGGAGCAGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Db	481	TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCAGAAGGAATCCCCCCAGCTGAG	540
Qу	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	600
Db	541	GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCT	600
Qу	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
Db .	601	ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC	660
Qу	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTCATCGTCTAC	720
Db	661	TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT	720
Qу	721	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	780
Db	721	GTGAACGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT	780
Qу	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
Db	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC	840
Qу	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC	900
Db	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGGAGC	900
Qу	901	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	960
Db	901	TGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGCCGC	960
Qу	961	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGGAGTGCCAGGGCACTGACCTGGAC	1020
Db	961	GAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTGGAC	1020
Qy	1021	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC	1080
Db	1021	ACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTGGCT	1080
Qy ,	1081	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATC	1140
Db	1081	CTCTACATCGGCCTTGTCGCTGTGGCTGTGCCTCTTCTTGCTGTTGCTGGCCCTTGGA	1140
Qу	1141	CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
Db	1141	CTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATCCTC	1200
Qу	1201	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
Db	1201	ACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTGCTC	1260
Qу	1261	ACCATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCCCGG	1320

--

	Db	1261	${\tt ACCATCCAGCCAGCACCACCACCTACCAGGGCAGTCTATGTTCGAGG}$	1320
	Qy	1321	CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1380
	Db	1321	CAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTGGGG	1380
	Qу	1381	GGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC	1440
	Db	1381	AGTGGCCGCCATACGTTGCACCACACCTCACCCACCTCTGAGGCTGAGGACTTCGTCTCC	1440
•	Qy	1441	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1500
	Db	1441	CGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCCTAC	1500
	Qу	1501	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1560
	Db	1501	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTCCTC	1560
	Qу	1561	ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
	Db	1561	ATACCCCGGATGCCATCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCACAAG	1620
	Qy	1621	CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC	1680
	Db	1621	CCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTTAGC	1680
	Qу	1681	TGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1740
	Db	1681	TGTGGGCCCCCAGGAGTCCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGTGGA	1740
	Qy	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
	Db	1741	GAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGTTGG	1800
	Qу		GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG	
	Db		GAGGATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAGCTGGAG	
	Qу		GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	
	Db		GCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGAGAGGCC	
	Qу		CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	
	Db		CTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCCTGTACG	
	Qу		TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	
	Db		TCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTCAAGGAG	
	Qy		GTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC	
	Db		GTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTCCTGCAC	
	Qy		TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG	
	Db	2101	TTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCCCTGTGG	2160

Qy Db		AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2220
		CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2280
Qy Db		
Db		CAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGCGACCTG 2280
Qy		GCCTGCAAGCTGTGGGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340
Db	2281	GCCTGCAAGGTGTGGGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATCAACTTC 2340
Qy	2341	AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400
Db	2341	AACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGGGGGGTC 2400
Qy	2401	CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460
Db	2401	
Qу	2461	TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
Db	2461	
Qу	2521	CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTC 2580
Db	2521	
Qу	2581	AACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC
Db	2581	AACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCAGCTGTG 2640
Qу	2641	GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
Db	2641	
RESULT 9 BD057524 LOCUS DEFINITION ACCESSION KEYWORDS SOURCE ORGANI REFERENC AUTHOR TITLE JOURNA COMMENT	ON N B B J S SM S E I S L P P	D057524 3014 bp DNA linear PAT 27-AUG-2002 etrin receptors. D057524 D057524.1 GI:22603130 P 2001505062-A/1. ynthetic construct ynthetic construct ther sequences; artificial sequences. (bases 1 to 3014) avigne, M.T., Leonardo, D.E., Hinck, L., Masu, M. and Masu, K.K. etrin receptors atent: JP 2001505062-A 1 17-APR-2001; HE REGENTS OF THE UNIV OF CALIFORNIA N JP 2001505062-A/1 D 17-APR-2001 F 19-FEB-1998 JP 1998536840 R 19-FEB-1997 US 08/808982

```
PΙ
                          MARC TESSIER LAVIGNE, DAVID E LEONARDO, LINDSAY HINCK, MASAYUKI
                  PΙ
                          MASU,
                          KAZUKO KEINO MASU
                  PΙ
                          C07K1/00,C07K14/00,C07K17/00,C07H21/02,C07H21/04,G01N33/53 CC
                  PC
                    Strandedness: Double;
                  CC
                          Topology: Linear;
                  FH
                          Key
                                                   Location/Qualifiers.
                                Location/Qualifiers
FEATURES
                                1. .3014
       source
                                /organism="synthetic construct"
                                /mol type="genomic DNA"
                                /db xref="taxon:32630"
ORIGIN
                                        83.5%;
                                                    Score 2252.2; DB 6;
                                                                                      Length 3014;
   Query Match
                                        89.7%;
                                                    Pred. No. 0;
   Best Local Similarity
   Matches 2419; Conservative
                                                   0; Mismatches 278;
                                                                                     Indels
                                                                                                     0; Gaps
                                                                                                                       0;
                  1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
Qу
                      1 ATGGCCGTCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60
Db
                 61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
Qу
                      61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120
Db
               121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
Qу
                      121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180
Db
               181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
Qу
                      181 GTGTTGTTGGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240
Db
               241 TGGGTGCGCCAGGTGACCACGTGATCGAGCGCACACAGACGGGAGCAGTGGGCTGCCC 300
Qу
                      Db
                301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
Qу
                      301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360
Db
               361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCAACAAGAGTCAGAAGGCC 420
Qу
                      361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420
Db
               421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
Qу
                                                 1111111 11 111
                421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480
Db
                481 TCCCTGGAGCAGGGCATCGTGCCGTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
Qу
                      11 | 11:11:11 | 1:11:1 | 1:11:1 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11:11 | 1:11 | 1:11:11 | 1:11:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1:11 | 1
                481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCAGAAGGAATCCCCCCAGCTGAG 540
Db
                541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
Qу
                      Db
```

Qy	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
Db	•		
Qу	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC	720
Db	661	TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT	720
Qу	721	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	780
Db	721	GTGAACGGTGGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT	780
Qу	781	GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
Db	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCCTTC	840
Qу	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC	900
Db	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGGAGC	900
Qу	901	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	960
Db	901	TGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGCCGC	960
Qу	961	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC	1020
Db	961	GAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTGGAC	1020
Qу	1021	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC	1080
Db	1021	ACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTGGCT	1080
Qу	1081	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGCTGCTGCTGCTGCTCATC	1140
Db	1081	$\tt CTCTACATCGGCCTTGTCGCTGTGGCTGTGCCTCTTCTTGCTGTTGCTGGCCCTTGGA$	1140
QУ		CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	
Db	1141	CTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATCCTC	1200
Qу	1201	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
Db	1201	ACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTGCTC	1260
Qу	1261	ACCATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCCCGG	1320
Db	1261	ACCATCCAGCCAGACCTCAGCACCACCACCTACCAGGGCAGTCTATGTTCGAGG	1320
Qу	1321	CAGGATGGGCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1380
Db	1321	CAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTGGGG	1380
Qу	1381	GGCGGCCGCCACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC	1440
Db	1381	AGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTCTCC	1440

			•
Qy	1441	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1500
Db	1441	CGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCCTAC	1500
Qу	1501	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1560
Db	1501	${\tt GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTCCTC}$	1560
Qу	1561	ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Db	1561	ATACCCCGGATGCCATCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCACAAG	1620
Qу	1621	CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC	1680
Db	1621	${\tt CCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTTAGC}$	1680
Qу	1681	TGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1740
Db	1681	TGTGGGCCCCAGGAGTCCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGTGGA	1740
Qу	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Db	1741	GAGCCCAGCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGTTGG	1800
Qу		GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG	1860
Db		GAGGATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAGCTGGAG	1860
Qу	1861	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	
Db	1861	GCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGAGAGGCC	
Qу	1921	CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	1980
Db	1921	$\tt CTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCCTGTACG$	1980
Qy	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	
Db	1981	TCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTCAAGGAG	
Qy	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC	2100
Db	2041	GTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTCCTGCAC	2100
Qy	2101	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG	2160
Db	2101	TTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCCCTGTGG	2160
Qу		AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG	2220
Db		AAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAACGGCACC	2220
Qy	2221	CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG	2280
Db	2221	CAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGCGACCTG	2280
Qv	2281	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2340

Db .	228	
Qy	234	AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400
Db	234	
Qy	240	CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460
Db	240	CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAGATCATC 2460
Qу	246	TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
Db	246	GCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAGAAACTT 2520
Qy	252	CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580
Db	252	CACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATGATCCTC 2580
Qy	258	AACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC
Db	258	AACCTATGGGAGGCACGCCCAACGGCAACCTCGGCCAGCTGGCAGCAGCTGTG 2640
Qy	264	GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
Db	264	GCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGTTGA 2697
RESULT 10 BC058084 LOCUS DEFINITION ACCESSION KEYWORDS SOURCE ORGANIS REFERENCE AUTHORS	ON I	3844 bp mRNA linear ROD 21-OCT-2003 Mus musculus unc-5 homolog A (C. elegans), mRNA (cDNA clone MGC:66671 IMAGE:6813463), complete cds. BCO58084 BCO58084.1 GI:34784158 MGC. Mus musculus (house mouse) Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. L (bases 1 to 3844) Btrausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,

Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

TITLE Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

MEDLINE 22388257 PUBMED 12477932

REFERENCE 2 (bases 1 to 3844)

AUTHORS Strausberg, R.
TITLE Direct Submission

JOURNAL Submitted (08-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov

COMMENT

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa cDNA Library Preparation: M. Bento Soares, University of Iowa cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 126 Row: b Column: 11

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 23346570.

FEATURES

Location/Qualifiers

source

1. .3844

/organism="Mus musculus"

/mol_type="mRNA"
/strain="C57BL/6"

/db xref="taxon:10090"

/clone="MGC:66671 IMAGE:6813463"

/tissue_type="Brain, mouse, 13.5, 14.5, 16.5, 17.5 dpc"

/clone lib="NIH BMAP FW0"

/lab host="DH10B"

/note="Vector: pYX-ASC"

gene

1. .3844

/gene="Unc5a"

/db_xref="LocusID:107448"

/db xref="MGI:894682"

CDS

252. .2780

/gene="Unc5a" /codon start=1

/product="Unc5a protein" /protein id="AAH58084.1"

```
/db xref="GI:34784159"
/db xref="LocusID:107448"
/db xref="MGI:894682"
translation="MAVRPGLWPALLGIVLTAWLRGSGAQQSATVANPVPGANPDLLP/
HFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTM
EVRINVSROOVEKVFGLEEYWCOCVAWSSSGTTKSOKAYIRIAYLRKNFEQEPLAKEV
SLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTPN
YTCVAKNIVARRRSASAAVIVYVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGE
ECRGADLDTRNCTSDLCLHTSSGPEDVALYIGLVAVAVCLILLLLVLVLIYCRKKEGL
DSDVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQDGPSPKF
QLSNGHLLSPLGSGRHTLHHSSPTSEAEDFVSRLSTQNYFRSLPRGTSNMAYGTFNFL
GGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPP
GVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEDVLHLGEESPSHLYYCQLEAG
ACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVACTSLEYNIRVYCLHDTHDALKE
VVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWN
GTQQYLHCTFTLERVNASTSDLACKVWVWQVEGDGQSFNINFNITKDTRFAEMLALES
EGGVPALVGPSAFKIPFLIRQKIITSLDPPCSRGADWRTLAQKLHLDSHLSFFASKPS
PTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC"
714. .926
/gene="Unc5a"
/note="IG; Region: Immunoglobulin"
/db xref="CDD:smart00409"
984. .1130
/gene="Unc5a"
/note="TSP1; Region: Thrombospondin type 1 repeats"
/db xref="CDD:smart00209"
1566. .1877
/gene="Unc5a"
/note="ZU5; Region: Domain present in ZO-1 and Unc5-like
netrin receptors"
/db xref="CDD:smart00218"
2496. .2684
/gene="Unc5a"
/note="DEATH; Region: DEATH domain, found in proteins
involved in cell death (apoptosis). Alpha-helical domain
present in a variety of proteins with apoptotic functions.
Some (but not all) of these domains form homotypic and
heterotypic dimers"
/db xref="CDD:smart00005"
                                                          ;
```

ORIGIN

misc feature

misc feature

misc feature

misc feature

Query Ma	72.6%;	Sc	ore	1957	.4;	DB 10	; Leng	th 3844	;			
Best Loc	cal S	Similarity	84.6%;	Pr	ed.	No.	0;					
Matches	2283	3; Conserva	itive	0;	Mis	smatcl	hes	246;	Indels	168;	Gaps	1;
Qy	1	ATGGCCGTCCG	GCCCGGC	CTGT	GGCC	CAGCG	CTCCI	rgggca	TAGTCCI	CGCCGCT	TGGCTC	60
		11111111111		$\Box\Box$	1111				1111111	1 + 1	11111	
Db	252	ATGGCCGTCCG	GCCCGGC	CTGT	GGCC	CAGCG	CTCCI	rgggca	TAGTCCT	CACTGCC	TGGCTT	311
Qy	61	CGCGGCTCGGG	TGCCCAG	CAGA	GTG	CCACC	GTGG	CCAACC	CAGTGCC	TGGTGCC	AACCCG	120
		11 11 11111	111111	1111	1111				1111111	1111111		
Db	312	CGTGGTTCGGG	TGCCCAG	CAGA	GTG	CCACA	GTGG	CCAACC	CAGTGCC	TGGTGCC	AACCCG	371
Qy	121	GACCTGCTTCC	CCACTTC	CTGG	TGG	AGCCC	GAGG/	ATGTGT	ACATCGI	CAAGAAC	AAGCCA	180
_			111111		1 11			1111	1111 11		11111	
Db	372	GACCTGCTGCC	CCACTTC	CTGG	TAG	AGCCG	GAGG	ACGTGT	ACATTGI	CAAGAAC	AAGCCC	431

Qy Db		GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG	
Qу		TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC	
Db	492	TGGGTTCGCCAGGTCGATCACGTCATTGAACGCAGCACTGACGGCAGCAGCAGCAGTTGCCA	551
Qу	301	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG	360
Db	552	ACCATGGAGGTCCGGATCAACGTATCAAGGCAGCAGGTCGAGAAAGTGTTTGGGCTGGAG	611
Qу	361	GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	420
Db	612	GAGTACTGGTGCCAGTGTGGGCATGGAGCTCCTCAGGAACCACCAAAAGCCAGAAGGCC	671
Qу	421	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
Db _.	672		731
Qу	481	TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Db	732	TCACTGGAGCAAGGCATTGTGCTACCTTGTCGCCCCCGGAAGGAA	791
Qy	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	600
Db	792		851
Qу	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
Db	852	ACACGGGAGCACAGCCTAGTCGTGCGGCAGGCCCGCCTGGCCGACACTCCCAACTACACC	911
Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTCATCGTCTAC	720
Db	912	TGCGTGGCCAAGAACATCGTGGCCCGTCGCCGAAGCGCCTCTGCGGCCGTCATTGTTTAT	971
Qy	721	GTGAACGGTGGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	780
Db	972	 GTG	974
Qy	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
Db	975		974
Qу	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC	900
Db	975		983
Qy	901	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	960
Db	984		1043
Qy	961	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC	1020
_			

10				
	Db	1104		1163
1164 CTCTACATCGGCCTCGTCGCCGTGGCCGTGTGCCTCATCTTGCTGCTGCTGCTCGTCCTCGTC 1223	Qу	1081		1140
	Db	1164		1223
1224 CTCATCTACTGCCGCAAGAAGGAAGGACTGGACTCAGCCTGCTCTT 1283	Qу	1141		1200
	Db	1224		1283
1284 ACCTCAGCCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGAATCCCCATCTGCTC 1343	Qу	1201		1260
	Db	1284		1343
Db	Qу	1261		1320
	Db	1344		1403
1404 CAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGGTCAGCCCATGGGC	Qy	1321		1380
	Db	1404	***************************************	1463
Db 1464 AGTGGCCGCCATACGCTGCACCACAGCTCCCCCACCTTGAGGCTGAGGACTTCGTCTCC 1523 Qy 1441 CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT 1500	Qy	1381		1440
	Db	1464		1523
Db	Qy	1441		1500
Db 1584 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAACACAGGAATCAGCCTCCTC 1643 Qy 1561 ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1620	Db	1524		1583
Db 1584 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAACAACAGGAATCAGCCTCCTC 1643 Qy 1561 ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1620	Qу	1501		1560
Db 1644 ATACCCCCGGACGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACTCTGCACAAG 1703 Qy 1621 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1680 Db 1704 CCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1763 Qy 1681 TGTGGACCCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATCGTTAGC 1763 Qy 1681 TGTGGACCCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1740 Db 1764 TGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCCATGGACCACTGCGG 1823 Qy 1741 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1800 Db 1824 GAGCCCAGCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCCTGTGAGGGCAGCTGG 1883 Qy 1801 GAGGATGTGCTGCACCTGGAGCCTTCCACCACCTCTACTACTGCCAGCTGGAG 1860 Db 1884 GAGGACGTGCTGCACCTGGGCGAGGAGGCCCCTCCCACCTCTACTACTGCCAGCTGGAG 1860 Db 1884 GAGGACGTGCTGCACCTTGGTGAGGAGCCCCTCCCACCTCTACTACTGCCAGCTGGAG 1860 Db 1884 GAGGACGTGCTGCACCTTGGTGAGGAGTCGCCCTCTACTACTACTGCCAGCTGGAG 1860 Qy 1861 GCCAGTGCCTGCACCTTGGTGAGGAGTCGCCCTCTACTACTACTGCCAGCTGGAG 1943	Db	1584		1643
Db 1644 ATACCCCCGGACGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACTCTGCACAAG 1703 Qy 1621 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1680	Qy	1561		1620
	Db	1644		1703
Db 1704 CCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCTATCGTTAGC 1763 Qy 1681 TGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1740	Qу	1621		1680
Db 1764 TGTGGGCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCCATGGACCACTGCGGG 1823 Qy 1741 GAGCCCAGCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1800 Db 1824 GAGCCCAGTCCCGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCCTGTGAGGGCAGCTGG 1883 Qy 1801 GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1860 Db 1884 GAGGACGTGCTGCACCTTGGTGAGGAGTCGCCCTCTACTACTACTGCCAGCTGGAG 1943 Qy 1861 GCCAGTGCCTGCACCTTGGTGAGGAGCCCCCTCTCATCTACTACTGCCAGCTGGAG 1943	Db .	1704		1763
Db 1764 TGTGGGCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCCATGGACCACTGCGGG 1823 Qy 1741 GAGCCCAGCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1800	QУ			1740
	Db			1823
Db 1824 GAGCCCAGTCCCGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCCTGTGAGGGCAGCTGG 1883 Qy 1801 GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1860	Qу	1741		1800
	Db	1824	GAGCCCAGTCCCGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCCTGTGAGGGCAGCTGG	1883
Db 1884 GAGGACGTGCACCTTGGTGAGGAGTCGCCCTCTCATCTCTACTACTGCCAGCTGGAG 1943 Qy 1861 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 1920	Qу	1801		1860
	Db	1884		1943
	Qу	1861		1920

Db	1944	GCCGGGGCCTGCTATGTCTTTACCGAGCAGCTAGGCCGCTTTGCCCTGGTGGGAGAGGCC	2003
Qу	1921	CTCAGCGTGGCTGCCCCAAGCCTCTCTGTTTGCGCCGGTGGCCTGCACC	1980
Db	2004	CTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCCCCTGTGGCCTGTACG	2063
Qу	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2040
Db	2064	TCCCTCGAGTACAACATCCGAGTGTACTGCCTGCACGACACCCACGATGCTCTCAAGGAG	2123
Qу	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC	2100
Db ·	2124	GTGGTGCAGCTGGAGAAGCAGCTGGTGGACAGCTGATCCAGGAGCCCCGTGTCCTGCAC	2183
Qу	2101	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG	
Db	2184	TTCAAAGACAGTTACCACAACCTACGTCTGTCCATCCACGACGTGCCCAGCTCCCTGTGG	
Qу	2161	AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG	2220
Db	2244	AAGAGCAAGCTCCTTGTCAGCTACCAGGAGATCCCTTTTTTACCACATCTGGAATGGCACT	2303
Qу	2221	CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG	2280
Db	2304	CAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCGTCAATGCCAGCACCAGCGACCTG	2363
Qу	2281	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2340
Db	2364	GCCTGCAAGGTGTGGGTGGCAGGTGGAGGGAGATGGACAGAGCTTCAACATCAACTTT	2423
Qу	2341	AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC	2400
Db	2424	AACATCACTAAGGACACGAGGTTTGCTGAAATGCTGGCTCTGGAGAGTGAAGGGGGGGTC	2483
Qу	2401	CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT	2460
Db	2484	CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAGATCATT	2543
Qy	2461	TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC	2520
Db	2544	ACCAGCCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGCGAACTCTAGCCCAGAAACTT	2603
Qу	2521	CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC	2580
Db		CACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATGATCCTC	
Qу	2581	AACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC	2640
Db	2664	AACCTATGGGAGGCGCGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCCGCAGCTGTG	2723
Qу		GCTGGACTGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 269	
Db	2724	GCCGGACTGGCCAGCCAGATGCTGGCCTCTTCACCGTGTCAGAGGCCGAGTGCTGA 278	30

```
BC009333
                                      2688 bp
                                                                   PRI 25-MAR-2004
LOCUS
                                                 mRNA
                                                          linear
            Homo sapiens unc-5 homolog A (C. elegans), mRNA (cDNA clone
DEFINITION
            IMAGE: 4126760), partial cds.
            BC009333
ACCESSION
            BC009333.2 GI:40226527
VERSION
KEYWORDS
SOURCE
            Homo sapiens (human)
  ORGANISM
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                (bases 1 to 2688)
REFERENCE
  AUTHORS
            Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
            Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
            Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
            Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
            Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
            Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
            Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
            Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
            Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
            McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
            Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
            Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
            Fahey, J., Helton, E., Ketteman, M., Madan, A., Rodrigues, S.,
            Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
            Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
            Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
            Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
            Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
  TITLE
            Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
            Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  JOURNAL
   PUBMED
            12477932
REFERENCE
                (bases 1 to 2688)
  AUTHORS
            Strausberg, R.
  TITLE
            Direct Submission
            Submitted (12-JUN-2001) National Institutes of Health, Mammalian
  JOURNAL
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
            NIH-MGC Project URL: http://mgc.nci.nih.gov
  REMARK
COMMENT
            On Dec 19, 2003 this sequence version replaced gi:14424611.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc mgc@nhgri.nih.gov
            Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
            Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
            Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
```

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL Plate: 26 Row: g Column: 22.

FEATURES

Location/Qualifiers

source

1. .2688

/organism="Homo sapiens"

/mol type="mRNA" /db xref="taxon:9606" /clone="IMAGE:4126760"

/tissue type="Brain, neuroblastoma"

/clone_lib="NIH_MGC_19" /lab host="DH10B-R" /note="Vector: pOTB7"

gene

<1. .2688 /gene="UNC5A"

/note="synonym: UNC5H1" /db_xref="LocusID:90249"

/db xref="MIM:607869"

CDS

<1. .1627 /gene="UNC5A" /codon start=2

/product="UNC5A protein" /protein id="AAH09333.2" /db xref="GI:40226528" /db xref="LocusID:90249" /db xref="MIM:607869"

/translation="DVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILT SGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPL GGGRHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGI SLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILA MDHCGEPSPDSWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGR FALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQ LIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFT LERVSPSTSDLACKLWVWOVEGDGOSFSINFNITKDTRFAELLALESEAGVPALVGPS AFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEA RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC"

misc feature 413. .724

/gene="UNC5A"

/note="ZU5; Region: Domain present in ZO-1 and Unc5-like

netrin receptors"

/db xref="CDD:smart00218"

misc feature

 $134\overline{3}$. .1594

/gene="UNC5A"

/note="DEATH; Region: DEATH domain, found in proteins involved in cell death (apoptosis). Alpha-helical domain present in a variety of proteins with apoptotic functions. Some (but not all) of these domains form homotypic and

heterotypic dimers"

/db xref="CDD:smart00005"

ORIGIN

Query Match 60.3%; Score 1625.4; DB 9; Length 2688; Best Local Similarity 99.9%; Pred. No. 4.1e-269;

	Matches	1626	5; Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
	Qy		GGACGTGGCCCTCTATG			111111	111111111	$\Pi\Pi$		
	Db .	1	GGACGTGGCCCTCTATG	TGGGC	CTCATCGCCGTGG	CCGTCI	GCCTGGTCC	TGCT	GCTGCT	60
	Qу	1131	TGTCCTCATCCTCGTTT							1190
	Db	61	TGTCCTCATCCTCGTTT.	ATTGC	CGGAAGAAGGAGG	GGCTGG	SACTCAGATG	TGGC	rgact _. c	120
	Qу	1191	GTCCATTCTCACCTCAG							1250
	Db	121	GTCCATTCTCACCTCAG							180
	Qу	1251	CCATCTGCTCACCATCC							1310
	Db	181	CCATCTGCTCACCATCC.							240
	Qy	1311	CTGTCCCCGGCAGGATG							1370
	Db	241	CTGTCCCCGGCAGGATG							300
	Qу	1371	CCCCTGGGTGGCGGCC							1430
	Db	301	CCCCTGGGTGGCGGCC							360
	Qу	1431	GTTCGTCTCCCGCCTCT							1490
	Db	361	GTTCGTCTCCCGCCTCT							420
	Qу	1491	CATGACCTATGGGACCT							1550
	Db	421	CATGACCTATGGGACCT							480
	Qу	1551	CAGCCTCCTCATCCCCC							1610
	Db	481	CAGCCTCCTCATCCCCC							540
	Qу		GCTGCACAAGCCGGAAG							
	Db		GCTGCACAAGCCGGAAG							
	Qу	1671	CATCGTTAGCTGTGGAC							1730
	Db	601	CATCGTTAGCTGTGGAC							660
	Qу	1731	CCACTGTGGGGAGCCCA							1790
	Db	661	CCACTGTGGGGAGCCCA							720
	Qy	1791	GGGCAGCTGGGAGGATG							1850
	Db	721	GGGCAGCTGGGAGGATG							780
	Qу	1851	CCAGCTGGAGGCCAGTG							1910
•	Db	781	CCAGCTGGAGGCCAGTG							840

.

	Qy	1911	GGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGT	1970
	Db	841	GGGAGAGGCCCTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGT	900
	Qy	1971	GGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2030
	Db .	901	GGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	960
	Qу	2031	ACTCAAGGAGGTGCTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACG	2090
	Db	961		1020
	Qy	2091	GGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	2150
	Db .	1021	GGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	1080
	Qy	2151	CTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTG	2210
	Db	1081	CTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTG	1140
	Qy .	2211	GAATGGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCAC	2270
	Db	1141	GAATGCCACCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCAC	1200
	Qy	2271	TAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGÁGCTTCAG	2330
	Db	1201	TAGTGACCTGCCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAG	1260
	Qу	2331	CATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGA	2390
	Db	1261	CATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGA	1320
	Qу	2391	AGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCA	2450
	Db	1321	AGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCA	1380
	Qу	2451	GAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGC	2510
	Db	1381	GAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGC	1440
•	Qу	2511	CCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGC	2570
	Db	1441	CCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGC	1500
	Qу	2571	CATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGC	2630
	Db	1501	CATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGC	1560
	Qу	2631	TGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGA	2690
•	Db	1561	TGCAGCAGTGGCTGGACTGGGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGA	1620
	Qу	2691	GTGCTGA 2697	
•	Db	1621	 GTGCTGA 1627	

```
RESULT 12
BD057525
LOCUS
          BD057525
                               1787 bp
                                                        PAT 27-AUG-2002
                                         DNA
                                                linear
DEFINITION
          Netrin receptors.
          BD057525
ACCESSION
          BD057525.1 GI:22603131
VERSION
          JP 2001505062-A/2.
KEYWORDS
SOURCE
          synthetic construct
 ORGANISM
          synthetic construct
          other sequences; artificial sequences.
REFERENCE
             (bases 1 to 1787)
 AUTHORS
          Lavigne, M.T., Leonardo, D.E., Hinck, L., Masu, M. and Masu, K.K.,
 TITLE
          Netrin receptors
          Patent: JP 2001505062-A 2 17-APR-2001;
 JOURNAL
          THE REGENTS OF THE UNIV OF CALIFORNIA
COMMENT
          PN
              JP 2001505062-A/2
          PD
              17-APR-2001
              19-FEB-1998 JP 1998536840
          PF
          PR
              19-FEB-1997 US
                              08/808982
              MARC TESSIER LAVIGNE, DAVID E LEONARDO, LINDSAY HINCK, MASAYUKI
          PT
          PΙ
              MASU,
          PΙ
              KAZUKO KEINO MASU
          PC
              C07K1/00,C07K14/00,C07K17/00,C07H21/02,C07H21/04,G01N33/53 CC
           Strandedness: Double;
          CC
              Topology: Linear;
          FH
                            Location/Qualifiers.
              Key
FEATURES
                  Location/Qualifiers
                  1. .1787
    source
                  /organism="synthetic construct"
                  /mol type="genomic DNA"
                  /db xref="taxon:32630"
ORIGIN
 Query Match
                      57.6%; Score 1552.4; DB 6;
                                                Length 1787;
 Best Local Similarity
                      98.5%; Pred. No. 1.5e-256;
 Matches 1651; Conservative
                            0; Mismatches
                                           16;
                                                                   8;
                                               Indels
                                                        9;
                                                            Gaps
       1025 GCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCT 1084
Qy
            1 GCAACTGTACCAGTGACCTCTG-GTACACACTGCTTCTGGCCCTGAGGACGTGGCCCTCT 59
Db
       1085 ATGTGGGCCTCATCGCCGTGGCCGTCTGCTGCTGCTGCTGCTGCTCATCCTCG 1144
Qy
            60 ATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCG 119
Db
Qу
       1145 TTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT 1204
            120 TTTATTGCCGGAAGAGGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT 179
Db
       1205 CAGGCTTCCAGCCCGTCAGCATC-AAGCCCAGCAAAGCAGACACCCCCATCTGCTCACC 1263
Qу
            Db
        180 CAGGCTTCCAGCCGTCAGCATCTAAGCCCAGCAAAGCAGCACCCCCATCTGCTCACC 239
       1264 ATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAG 1323
Qу
            Db
        240 ATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAG 299
```

Qу		GATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGC	
Db		GATGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGC GGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC	
Qy Db		GGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC	
Qy		CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGG	
Db			
Qy		ACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATC	
Db			
Qy	1564	CCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG	1623
Db ·	540		599
Qy	1624	GAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGT	1683
Db	600	GAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGT	659
Qy	1684	GGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAG	1743
Db	660	GGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAG	719
Qу	1744	CCCAGCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAG	1803
Db	720	CCCAGCCCTGACAGCTGGAGCCTGGCCCTCAAAAAGCAGTCGTGCGAGGG-AGCTGGGAG	778
Qy	1804	GATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCC	1863
Db	779	GATGT-CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCC	837
Qу	1864	AGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTC	1923
Db	838	AGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTC	897
Qу	1924	AGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCC	1983
Db	898	AGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCC	957
QУ	1984	CTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTG	2043
Db	958	CTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTG	1017
Qy	2044	GTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTC	2103
Db	1018	GTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTT-	1076
Qy	2104	AAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAG	2163
Db	1077	AAGGACAGTTACCACAACCTGCCCTATCATCCACGATGTGCCCAGCTCCCTGTGGAAG	1134

•

Qy		4 AGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAG 2223						
Db	113	35 AGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAG 1194						
Qy	222	24 CGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCC 2283						
Db	119	25 CGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCC 1254						
Qу	228	4 TGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAAC 2343						
Db .	125	5 TGCAAGCTGTGGGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAAC 1314						
Qу	234	4 ATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCA 2403						
Db	131	.5 ATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCA 1374						
QУ	240	4 GCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCC 2463						
Db	137	5 GCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCC 1434						
Qу	246	4 AGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCAC 2523						
Db	143	5 AGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCAC 1494						
Qу	252	4 CTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAAC 2583						
Db	149	5 CTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAAC 1554						
Qу	258	4 CTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC						
Db	155	55 CTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC						
Qу	264	4 GGACTGGGCCAGCCAGACGCTGGCCTC-TTCACAGTG-TCGGAGGCTGAGTGCTGA 2697						
Db	161							
RESULT 1	3							
AX054976 LOCUS		AX054976 9700 bp DNA linear PAT 13-JAN-2001						
DEFINITI	ON	Sequence 91 from Patent WO0073328.						
ACCESSION VERSION	N	AX054976 AX054976.1 GI:12228344						
KEYWORDS								
SOURCE	~~.	synthetic construct						
ORGANI	SM	synthetic construct other sequences; artificial sequences.						
REFERENC	E	1						
AUTHORS		van Criekinge, W., Roelens, I., Bogaert, T. and Verwaerde, P.						
TITLE JOURNAL		Unc-5 constructs and screening methods Patent: WO 0073328-A 91 07-DEC-2000;						
DE AUTODO		Devgen NV (BE)						
FEATURES sou		Location/Qualifiers 19700						
554		/organism="synthetic construct"						
		/mol_type="unassigned DNA"						
		/db_xref="taxon:32630"						

ORIGIN

		48.3%; Score 1302.8; DB 6; Length 9700; Similarity 99.8%; Pred. No. 8.4e-214; 4; Conservative 0; Mismatches 2; Indels 0; Gaps	0;
Qу	1392	CACACTGCACCACGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCAC	1451
Db	7		66
Qу	1452	CCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAA	1511
Db	67		126
Qу	1512	CTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGA	1571
Db	127		186
Qу	1572	TGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGT	1631
Db	187	TGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGT	246
Qу	1632	GAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCC	1691
Db	247	GAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCC	306
Qу	1692	TGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCC	1751
Db	307	TGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCC	366
Qу	1752	TGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCT	1811
Db	367	TGACAGCTGGAGCCTCCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCT	426
Qу	1812	GCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTG	1871
Db	427	GCACCTGGGCGAGGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTG	486
Qу	1872	CTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGC	1931
Db	487	CTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGC	546
Qу	1932	TGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC	1991
Db	547	TGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC	606
Qу	1992	CAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCT	2051
Db	607	CAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCT	666
Qу	2052	GGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAG	2111
Db	667	GGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAG	726
Qу	2112	TTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCT	2171

Db	727	TTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCT 786						
Qу	2172	CCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTT 2231						
Db	787	CCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTT 846						
Qу	2232	GCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCT 2291						
Db	847	GCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCT 906						
Qу	2292	GTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAA 2351						
Db	907	GTGGGTGTGGCAGGTGGAGGGCGACGGCAGAGCTTCAGCATCAACTTCAACATCACCAA 966						
Qy	2352	GGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGT 2411						
Db	967	GGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCAAGCCCTGGT 1026						
Qy	2412	GGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGA 2471						
Db	1027	GGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGA 1086						
Qy	2472	CCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAG 2531						
Db	1087	CCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAG 1146						
Qу	2532	CCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGA 2591						
Db	1147	CCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGA 1206						
Qу	2592	GGCGCGCCACTTCCCCAACGCCAACCTCAGCCAGCTGGCTG						
Db	1207	GGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTG						
Qу	2652	CCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697						
Db	1267	CCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 1312						
RESULT 1	.4							
MMU72634 LOCUS		MU72634 9299 bp mRNA linear ROD 13-MAY-1997						
DEFINITI	ON M	Mus musculus rostral cerebellar malformation protein (rcm) mRNA, complete cds.						
ACCESSIC VERSION	ON U	172634 172634.1 GI:2088526						
KEYWORDS	_	7200111 011200020						
SOURCE		(us musculus (house mouse)						
ORGANI	E	Mus musculus Cukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
REFERENC	CE 1	ammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. (bases 1 to 9299)						
AUTHOR		Ackerman, S.L., Kozak, L.P., Przyborski, S.A., Rund, L.A., Boyer, B.B.						
TITLE	J	and Knowles, B.B. The mouse rostral cerebellar malformation gene encodes an UNC-5-like protein						
JOURNA								

```
97271898
 MEDLINE
  PUBMED
           9126743
REFERENCE
              (bases 1 to 9299)
           Ackerman, S.L., Kozak, L.P., Rund, L.A. and Knowles, B.B.
 AUTHORS
 TITLE
           Direct Submission
           Submitted (25-SEP-1996) The Jackson Laboratory, 600 Main Street,
 JOURNAL
           Bar Harbor, ME 04609, USA
FEATURES
                   Location/Qualifiers
                   1. .9299
    source
                   /organism="Mus musculus"
                   /mol type="mRNA"
                   /strain="C57B6/SJL"
                   /db xref="taxon:10090"
                   /chromosome="3"
                   /map="between Amy1 and Mpmv9"
                   1. .9299
    gene
                   /gene="rcm".
                   151. .2946
    CDS
                   /gene="rcm"
                   /function="transmembrane receptor"
                   /note="similar to C. elegans UNC-5"
                   /codon start=1
                   /product="rostral cerebellar malformation protein"
                   /protein id="AAB54103.1"
                   /db xref="GI:2088527"
                   translation="MRKGLRATAARCGLGLGYLLQMLVLPALALLSASGTGSAAQDDE/
                   FFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQ
                   KDHVVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYV
                   RIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEDIIDPAEDRNFYI
                   TIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRC
                   GRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIACTTLCPVDGRWTSWSKWSTCGTECTH
                   WRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPDSDDVALYVGIVIAVTVCLA
                   ITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQDLLAVPPDLTSAAAMYR
                   GPVYALHDVSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFSSKLSPQMTQS
                   LLENEALNLKNQSLARQTDPSCTAFGTFNSLGGHLIIPNSGVSLLIPAGAIPQGRVYE
                   MYVTVHRKENMRPPMEDSQTLLTPVVSCGPPGALLTRPVILTLHHCADPSTEDWKIQL
                   KNQAVQGQWEDVVVVGEENFTTPCYIQLDAEACHILTENLSTYALVGQSTTKAAAKRL
                   KLAIFGPLCCSSLEYSIRVYCLDDTODALKEVLOLEROMGGOLLEEPKALHFKGSIHN
                   LRLSIHDIAHSLWKSKLLAKYQEIPFYHIWSGSQRNLHCTFTLERLSLNTVELVCKLC
                   VRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPASTITTVTGPSAFSIPLPIRQKLCSSL
                   DAPQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLE
                   EMGRHETVVSLAAEGOY"
ORIGIN
 Query Match
                        36.8%; Score 992; DB 10; Length 9299;
 Best Local Similarity
                        62.7%; Pred. No. 1.9e-160;
 Matches 1660; Conservative 0; Mismatches 925; Indels
                                                           63; Gaps
         104 TGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACA 163
Qу
                                 | \cdot |
                        308 TTCCATCTGACCCACCTGAGCCATTGCCACACTTCCTCATTGAGCCCGAGGAAGCTTACA 367
Db
         164 TCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCT 223
Qу
             Db
```

224 TCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACG 283

Qу

Db	428	TCAAGTGCAACAGCGAGTGGGTTCATCAGAAGGACCACGTAGTAGACGAGAGAGTAGATG	487
Qу	284	GGAGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGA	343
Db	488	AAACCTCTGGTCTAATTGTGAGAGAGAGTGAGCATTGAGATTTCACGCCAGCAGGTGGAGG	547
Qy	344	AGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCA	403
Db	548	AACTGTTTGGGCCTGAAGATTACTGGTGCCAGTGTGTGGCCTGGAGCTCAGCAGGCACTA	607
Qу	404	CCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGC	463
Db	608	CGAAGAGTCGGAAGGCATACGTGCGCATTGCGTATCTGCGGAAGACATTCGAGCAGGAAC	667
Qy	464	CGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGG	523
Db	668	CCTTGGGAAAGGAAGTGTCCTTGGAGCAGGAAGTCTTACTCCAGTGTCGGCCACCTGAAG	727
Qу	524	GCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGG	583
Db	728	GGATCCCAGTGGCTGAGGTGGAATGGCTAAAGAATGAAGACATAATTGATCCTGCTGAAG	787
Qy	584	ACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTG	643
Db	788	ATCGGAACTTTTATATTACTATCGATCACAACCTGATCATCAAGCAAG	847
Qy	644	ACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCG	703
Db	848	ATACAGCAAATTATACCTGTGTTGCCAAAAATATTGTTGCCAAGAGAAAAAGCACCACAG	907
Qy	704	CTGCTGTCATCGTCAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCA	763
Db	908	CCACTGTCATCGTGTATGTTAATGGTGGCTGGTCCACCTGGACAGAGTGGTCTGTGTA	967
Qy	764	GCGCCAGCTGTGGGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTC	823
Db .	968	ACAGCCGCTGTGGGCGAGGATATCAGAAACGCACAAGAACCTGCACCAACCCAGCCCCAC	1027
Qу	824	TCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGT	883
Db .	1028	TCAATGGTGGGGCCTTCTGTGAGGGGCAGAGTGTGCAGAAAATAGCATGCACTACGTTAT	1087
Qу	884	GCCCAGTAGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCA	943
Db	1088	GTCCAGTGGATGGTAGGTGGACTTCATGGAGCAAATGGTCAACCTGTGGGACTGAATGCA	1147
Qу	944	CCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCC	1003
Db	1148	CCCACTGGCGCAGGAGGGGGTAACAGCACCAGCCCCCAAGAACGGGGGTAAGGACTGTG	1207
Qу	1004	AGGGCACTGACCTGGACACCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTG	1063
Db	1208	ATGGCCTGGTCCTCCAATCCAAGAACTGCACTGATGGGCTGTGCATGCA	1267
Qу	1064	GCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCC	1120

		·	
Db ·	1268	ACTCAGATGATGTGGCTCTCTACGTGGGGATTGTGATCGCTGTAACAGTCTGTCT	1327
Qу	1121	TGCTGCTGCTCGTCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATG	1180
Db	1328	TCACTGTTGTGGGGCCCTGTTTGTGTATCGGAAGAACCACCGTGACTTTGAGTCTGACA	1387
Qу	1181	TGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAG	1240
Db	1388		1438
Qу	1241	CAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCACCACCACCACCACCACC	1300
Db	1439		1498
Qy	1301	AGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAA	1355
Db	1499	GGGGACCTGTCTATGCTCTGCATGATGTCTCAGACAAAATCCCAATGACCAACTCTCCAA	1558
Qy	1356	TGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCAC	1415
Db	1559	TTCTGGACCCACTACCCAACTTGAAAATCAAAGTGTACAACAGCTCAGGTGCTGTCACTC	1618
Qу	1416	CTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCC	1453
Db	1619	CTCAGGATGACCTTGCCGAGTTCTCATCCAAACTGTCACCCCAGATGACCCAGTCCTTGC	1678
Qy	1454	AGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCA	1489
Db	1679	TAGAGAATGAGGCCCTTAACCTGAAGAACCAGAGCCTCGCAAGACAGAC	1738
Qy	1490	ACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTA	1549
Db	1739	GCACAGCATTTGGTACCTTCAACTCTCTTGGGGGTCACCTCATCATTCCTAATTCAGGAG	1798
Qy	1550	TCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCA	1609
Db	1799	TAAGCTTGCTGATTCCCGCTGGGGCCATTCCTCAGGGGAGAGTCTATGAAATGTATGT	1858
Qy	1610	CGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTG	1669
Db	1859	CTGTACACAGGAAAAATATGAGGCCCCCCATGGAAGACTCTCAGACCCTACTTACCC	1918
Qу	1670	CCATCGTTAGCTGTGGACCCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGG	1729
Db	1919	CTGTGGTGAGCTGTGGGCCTCCTGGAGCTCTGACCCGCCCTGTCATCCTCACTCTGC	1978
Qy	1730	ACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCG	1789
Db	1979	ATCACTGTGCAGACCCCAGCACCGAGGACTGGAAGATCCAGCTCAAAAACCAGGCAGTGC	2038
Qy	1790	AGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACT	1849
Db	2039	AGGGACAATGGGAGGATGTTGTGGTGGTTGGGGAGGAGAACTTCACAACCCCCTGTTACA	2098
Qy	1850	GCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGG	1909
Db	2099	TTCAGCTGGATGCAGAGGCTTGCCATATCCTCACAGAGAACCTCAGTACCTATGCCCTGG	2158

.

Qу		TGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGG	
Db	2159	TTGGGCAGTCCACCACAAGCAGCTGCCAAGCGTCTTAAACTGGCCATCTTTGGGCCCC	2218
Qy	1970	TGGCCTGCACCTCCAGTACAACATCCGGGTCTACTGCCTGC	2029
Db	2219	${\tt TCTGCTGCTCTTCCCTGGAGTACAGCATTAGAGTCTACTGCCTGGATGACACACAGGATG}$	2278
Qу	2030	CACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCAC	2089
Db	2279	CCCTGAAGGAAGTTCTACAACTGGAGAGGCAAATGGGAGGACAGCTCCTAGAAGAACCCA	2338
Qy	2090	GGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	2149
Db.	2339	AGGCTCTTCATTTTAAAGGCAGCATCCACAACCTGCGCCTGTCTATTCATGACATCGCCC	2398
Qу	2150	GCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCT	2209
Db	2399		2458
Qу	2210	GGAATGGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCA	2269
Db	2459	GGAGTGGCTCTCAAAGAAACCTCCACTGCACCTTCACTCTGGAAAGACTCAGCCTAAACA	2518
Qy	2270	$\tt CTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCA$	2329
Db	2519	CAGTGGAACTGGTTTGCAAACTCTGTGTGCGGCAGGTTGAAGGAGAAGGGCAGATCTTCC	2578
Qу	2330	GCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTG	2389
Db	2579		2638
Qy	2390	AAGCGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGC	2449
Db	2639		2698
Qy	2450	AGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGG	2509
Db	2699	AGAAGCTATGCAGCAGCCTGGATGCCCCTCAAACAAGAGGCCATGACTGGAGGATGCTGG	2758
Qy	2510	CCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAG	2569
Db	2759		2818
Qу	2570	CCATGATCCTCAACCTGTGGGAGGCGCGCACCTTCCCCAACGGCAACCTCAGCCAGC	2629
Db	2819		2878
Qу	2630	CTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTG	2689
Db	2879		2938
Qу	2690	AGTGCTGA 2697	
Db	2939	AGTATTGA 2946	

```
RESULT 15
AY187310
LOCUS
```

AY187310 2962 bp mRNA linear VRT 06-JUN-2003

DEFINITION Gallus gallus UNC5-like protein 3 mRNA, complete cds.

ACCESSION AY187310

VERSION AY187310.1 GI:31442350

KEYWORDS

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 2962)
AUTHORS Guan, W. and Condic, M.L.

TITLE Characterization of Netrin-1, Neogenin and cUNC-5H3 expression

during chick dorsal root ganglia development

JOURNAL Gene Expr. Patterns 3, 369-373 (2003)

REFERENCE 2 (bases 1 to 2962)
AUTHORS Guan, W. and Condic, M.L.
TITLE Direct Submission

JOURNAL Submitted (26-NOV-2002) Neurobiology & Anatomy, University of Utah,

20 North, 1900 East, Salt Lake City, UT 84132-3401, USA

FEATURES Location/Qualifiers

source 1. .2962

/organism="Gallus gallus"

/mol_type="mRNA"
/db xref="taxon:9031"

CDS 1. .2796

/codon start=1

/product="UNC5-like protein 3"

/protein_id="AAO67275.1" /db xref="GI:31442351"

translation="MGKGLEGTAARCGLGMGYLLHSVVLPALAVLGASRPGSAAQDDD/ FFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQ KDHVVDERVDETSGLIVCEVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYV RIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEEVIDPVEDRNFYI TIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSACNSRC GRGFQKRTRTCTNPAPLNGGAFCEGQNVQKIACTTLCPVDGKWTSWSKWSTCGTECTH WRRRECTAPAPKNGGKDCEGLVLQSKNCTDGLCMQAAPDSDDVALYVGIVIAVIVCLA ISVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQDLLAVPPDLTSAAAMYR GPVYALHDVSDKIPMTNSPILDPLPNLKIKVYNTSGAVTPQDELSDFSSKLSPQITQS LLENETLNVKNQSLARQTDPSCTAFGTFNSLGGHLVIPNSGVSLLIPAGAVPQGRVYE ${ t MYVTVHRKEGMRPPVEDSQTLLTPVVSCGPPGALLTRPVVLTMHHCAEPNMDDWQIQL$ KHQAGQGPWEDVVVVGEENFTTPCYIQLDPEACHILTETLSTYALVGQSITKAAAKRL KLAIFGPLSCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEEPKTLHFKGSTHN LRLSIHDIAHSLWKSKLPAKYQEIPFYHIWSGCQRNLHCTFTLERFSLNTLELVCKLC VRQVEGEGQIFQLNCSVSEEPTGIDYPIMDSAGSITTIVGPNAFSIPLPIRQKLCSSL DAPQTRGHDWRMLAHKLKLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLE EMGRHETVVSLAAEGNY"

ORIGIN

Query Match 36.7%; Score 988.8; DB 5; Length 2962;

Best Local Similarity 62.6%; Pred. No. 8e-160;

Matches 1658; Conservative 0; Mismatches 927; Indels 63; Gaps 5;

•	Db	158	TTCCTTCTGATCCTCCAGAGCCATTGCCCCACTTTCTCATTGAACCCGAAGAAGCTTACA	217
	Qy Db		TCGTCAAGAACAAGCCAGTGCTGCTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCT	
	Qу		TCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACG	
	Db .	-		
	Qy	284	GGAGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGA	343
	Db	338		397
	QУ	344	AGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCA	403
	Db	398	AGCTCTTTGGACCCGAGGACTACTGGTGCCAGTGTGTCGCCTGGAGCTCAGCTGGCACCA	457
	Qу	404	CCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGC	463
	Db	458	CCAAGAGCCGCAAGGCCTACGTCCGCATTGCATATCTCAGAAAGACTTTTGAGCAGGAGC	517
	Qy	464	CGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGG	523
	Db	518	CGCTGGGGAAAGAAGTGTCCCTGGAGCAAGAGGTCCTGCTCCAGTGCCGTCCTCCTGAAG	577
	Qy		GCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGG	
	Db	578	GCATTCCAGTAGCTGAGGTAGAGTGGCTGAAGAATGAAGAGGTGATCGATC	637
	Qу		ACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTG	
	Db		ACCGAAATTTTTACATCACCATTGATCACAACCTGATCATCAAGCAAG	
	Qу		ACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCG	
	Db		ACACGGCTAACTACACCTGTGTTGCCAAAAACATTGTGGCCAAAAGGAAAAGCACGACAG	
	Qy		CTGCTGTCATCGTCTACGTGAACGGTGGGTCGACGTGGACCGAGTGGTCCGTCTGCA	
÷	Db		CAACTGTGATTGTCTATGTGAATGGAGGCTGGTCTACCTGGACCGAGTGGTCAGCGTGCA	
	Qy		GCGCCAGCTGTGGGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTC	
	Db		ACAGCCGCTGTGGGAGAGGCTTCCAGAAGCGCACAAGGACCTGCACTAACCCTGCCCCAC	
	Qy .		TCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGT	
	Db		TCAATGGGGGGCCTTCTGCGAGGGGCAAAATGTTCAGAAAATAGCTTGCACCACCCTGT	
	ДУ		GCCCAGTAGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCA	
	Db		GTCCAGTGGATGGCAAATGGACGTCCTGGAGCAAGTGGTCCACTTGTGGCACAGAGTGTA	
	Qy	944	CCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCC	1003

	Db	998	$\tt CCCACTGGCGCCGGAGGGGGGGGGGGGGGGGGGGGGGGG$	1057
	Qy	1004	AGGGCACTGACCTGGACACCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTG	1063
	Db	1058		1117
	Qy	1064	GCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCC	1120
	Db	1118	ACTCGGATGATGTTGCTCTACGTGGGGATTGTCATTGCTGTGATTGTGTGCCTGGCTA	1177
	Qy	1121	TGCTGCTGCTCGTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATG	1180
	Db	1178	TTTCTGTGGTTGTGGCCCTGTTTGTCTATCGCAAGAACCACCGTGACTTTGAGTCAGATA	1237
	Qy	1181	TGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAG	1240
	Db	1238	TTATCGACTCATCGGCGCTAAATGGGGGATTTCAGCCTGTTAACATCAAGGCTGCAAGAC	1297
	Qy	1241	CAGACACCCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACC	1300
	Db	1298	AAGACCTCTTGGCAGTGCCACCAGACCTCACTTCTGCTGCAGCCATGTACA	1348
	Qу	1301	AGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAA	1355
	Db	1349	GGGGGCCTGTGTATGCCTTGCATGATGTCTCTGATAAAATCCCAATGACCAATTCTCCGA	1408
	Qу	1356	TGGGCACCTGCTCAGCCCCTGGGTGGCGCCCACACACTGCACCACAGCTCTC	1411
	Db	1409	TCCTGGACCCACTGCCCAATCTGAAGATTAAAGTTTATAACACCTCTGGAGCAGTCACCC	1468
	Qy	1412	CCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCC	1453
	Db	1469	CCCAGGATGAACTCTCTGACTTCTCCTCCAAGCTGTCCCCACAGATTACCCAGTCTCTGT	1528
	Qу	1454	AGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCA	1489
	Db	1529	TGGAGAATGAGACTCTGAACGTGAAGAACCAAAGCCTTGCACGGCAAACAGACCCATCCT	1588
	Qy	1490	ACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTA	1549
	Db	1589	GCACTGCATTTGGGACCTTCAACTCGTTAGGGGGCCACCTAGTAATTCCTAATTCAGGAG	1648
. •	Qy	1550	TCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCA	1609
	Db	1649	TGAGCTTGCTGATCCCAGCAGGGGCTGTTCCCCAAGGAAGAGTCTATGAAATGTATGT	1708
	Qy	1610	CGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTG	1669
	Db	1709	CAGTCCACAGGAAGGAGGCATGAGACCACCTGTAGAAGACAGCCAGACGCTGCTGACAC	1768
	Qy	1670	CCATCGTTAGCTGTGGACCCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGG	1729
	Db	1769	CAGTGGTGAGCTGTGGCCCACCAGGAGCGCTGCTGACCCGACCCGTTGTGCTGACCATGC	1828
	Qy	1730	ACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCG	1789
	Db	1829	ACCACTGTGCTGAGCCCAACATGGATGACTGGCAGATCCAGCTCAAGCACCAGGCAGG	1888

Qу	1790	AGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACT	1849
Db	1889	AGGGACCATGGGAGGATGTAGTGGTGGTCGGGGAGGAAAACTTCACCACTCCATGCTACA	1948
Qу	1850	GCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGG	1909
Db	1949	TCCAGCTGGACCCAGAGGCCTGTCATATCCTGACGGAGACCCTCAGCACGTACGCCTTGG	2008
Qу	1910	TGGGAGAGGCCCTCAGCGTGCCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGG	1969
Db	2009	TGGGACAATCCATCACCAAAGCAGCCAAACGTCTCAAATTGGCCATCTTTGGACCAC	2068
Qу	1970	TGGCCTGCACCTCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2029
Db	2069	TGTCCTGTTCCTCACTGGAGTACAGCATCCGCGTCTACTGCCTCGATGACACACAGGATG	2128
Qy	2030	CACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCAC	2089
Db	2129	CCCTGAAGGAGGTCCTCCAGCTTGAGCGGCAGATGGGTGGG	2188
Qу	2090	GGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	2149
Db	2189	AAACTTTGCATTTTAAGGGAAGTACCCACAACCTGCGCTTATCCATTCATGACATTGCCC	2248
Qy	2150	GCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCT	2209
Db	2249	ACTCTCTCTGGAAGAGCAAACTGCCGGCTAAATACCAGGAGATTCCTTTTTACCACATCT	2308
Qу	2210	GGAATGCCACCAGCAGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCA	2269
Db	2309	GGAGTGGGTGCCAGAGGAACTTGCACTGCACCTTCACGCTGGAACGATTCAGTCTCAATA	2368
Qу	2270	CTAGTGACCTGCCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGCAGAGCTTCA	2329
Db	2369	CCCTGGAGCTCGTCTGCAAACTCTGTGTGCGGCAAGTCGAAGGAGAAGGGCAGATCTTCC	2428
Qу	2330	GCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTG	2389
Db	2429	AGCTGAACTGCTCAGTATCAGAGGAACCCACTGGCATTGATTATCCCATCATGGATTCAG	2488
QУ	2390	AAGCGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGC	2449
Db	2489	CAGGCAGCATCACAACGATAGTTGGGCCCAACGCTTTCAGCATCCCCCTCCCAATAAGGC	2548
Qу	2450	AGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGG	2509
Db	2549	AGAAGCTCTGCAGCAGCCTGGATGCACCCCAGACCCGGGGCCATGACTGGAGGATGCTGG	2608
Qy	2510	CCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAG	2569
Db	2609	CCCACAAGCTGAAATTGGACAGGTACCTAAATTATTTTGCTACGAAGTCGAGTCCCACTG	2668
Qу	2570	CCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGC	2629
Db	2669	GGGTGATCCTGGATCTCTGGGAAGCCCAGAATTTCCCTGATGGCAACCTGAGCATGCTGG	2728

Qу	2630	CTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTG 2689	
Db	2729	CAGCAGTGCTGGAAGAAATGGGACGACATGAAACCGTTGTTTCTTTGGCAGCAGAAGGAA 2788	
Ov	2690	AGTGCTGA 2697	
ΔÃ			
Db	2789	ATTACTGA 2796	

Search completed: March 6, 2005, 05:25:12 Job time: 11476.2 secs

GenCore version 5.1.6 Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2005, 14:42:51; Search time 1373 Seconds

(without alignments)

11628.216 Million cell updates/sec

Title: US-10-624-932-1_COPY_46_2742

Perfect score: 2697

Sequence: 1 atggccgtccggcccggcct.....tgtcggaggctgagtgctga 2697

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: N_Geneseq_16Dec04:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002as:*

7: geneseqn2002bs:*

8: geneseqn2003as:*

9: geneseqn2003bs:*

10: geneseqn2003cs:*

11: geneseqn2003ds:*

12: geneseqn2004as:*

13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2697	100.0	2752	6	ABK37922	Abk37922 cDNA enco
2	2697	100.0	2752	12	ADH71617	Adh71617 Human gen
3	2687.4	99.6	2697	6	ABK52891	Abk52891 Human net
4	2621.4	97.2	2881	6	ABK49422	Abk49422 DNA encod

```
5 2621.4
                    97.2
                               2881 12 ADH71609
                                                                                         Adh71609 Human gen
                               2880 12 ADH71633
 6 2619.8
                    97.1
                                                                                         Adh71633 Human gen
                            2881 12 ADH71649
 7 2619.8
                    97.1
                                                                                         Adh71649 Human gen
                                                                     Adh71635 Human gen
Adh71637 Human gen
Adh71641 Human gen
Adh71631 Human gen
Adh71645 Human gen
Adh71645 Human gen
Adh71645 Human gen
Adh71647 Human gen
Adh71643 Human gen
Adh71643 Human gen
Adh71643 Human gen
Adh71644 Human gen
Adh71647 Human gen
Adh71647 Human gen
Adh71649 Human man
Adh71649 Human man
Adh71649 Human REP
Adh71621 Human gen
Adh71623 Human gen
Adh71623 Human gen
Adh71623 Human gen
Adh71624 Human gen
Adh71629 Human LEC
Aav52940 Rat UNC-5
Adn95100 Human LEC
Aav52941 Human UNC
Aac90958 Plasmid p
Aah99530 Human pro
Adh71619 Human gen
Adh71619 Human gen
Adh71613 Human gen
Adh71611 Human gen
Adh71611 Human gen
Adh71611 Human gen
Abt06279 Human NOV
Abt06280 Human NOV
Abt06280 Human NOV
Abd93898 Human tra
Adr99249 Human 1RO
Adr07892 Full leng
Aas21316 Human cDN
Aca03675 cDNA enco
Abx89213 DNA encod
Acd41867 Human sec
 8 2619.8
                    97.1 2881 12 ADH71635
                                                                                         Adh71635 Human gen
 9 2619.8
                    97.1 2881 12 ADH71637
                                                                                         Adh71637 Human gen
10 2619.8
                    97.1 2881 12 ADH71641
11 2619.8
                    97.1 2881 12 ADH71629
                    97.1 2881 12 ADH71631
12 2619.8
                   97.1 2881 12 ADH71645

97.1 2881 12 ADH71627

97.1 2881 12 ADH71639

97.1 2881 12 ADH71643

97.1 2881 12 ADH71645

97.1 2881 12 ADH71625

97.1 2881 12 ADH71647

97.1 2881 10 ADG42568

87.8 2907 4 AAK52261
13 2619.8
14 2619.8
15 2619.8
16 2619.8
17 2619.8
18 2619.8
19 2618.2
20 2368.6
                               3561 12 ADL06497
21
         2343
                    86.9
                    86.9 3580 6 ABK15169
22
         2343
                    85.9 2575 12 ADH71621
         2318
23
                    84.0 2463 12 ADH71623
24 2266.2
25 2252.2 83.5 2697 6 AAS16843
                    83.5 3014 2 AAV52940
26 2252.2
                             2635 11 ADN95100
27 1625.4
                    60.3
                            1787 2 AAV52941
28 1552.4
                    57.6
29 1302.8
                            9700 4 AAC90958
                  48.3
                   44.5 1321 4 AAH99530
30
    1200.6
31
       987.6
                    36.6 1002 12 ADH71619
32
       985.6
                    36.5 1009 12 ADH71613
                               3646 5 AAS75738
33
           970
                    36.0
                               3646 13 ADR99257
           970
34
                    36.0
                               2796 10 AAL56266
35
       968.8
                    35.9
                               993 12 ADH71611
36
       943.4
                    35.0
37
       936.2
                               2860 6 ABT06279
                    34.7
38
       936.2
                    34.7 2860 6 ABT06280
39
       912.2
                    33.8 2895 6 ABQ93898
                    33.5 2986 13 ADR99249
40
       904.2
                    33.5 3933 13 ADR07892
41
       904.2
                    33.4 3884 4 AAS21316
42
           901
                               3884 8 ACA03675
43
           901
                    33.4
                               3884 8 ABX89213
44
           901
                    33.4
                               3884 8 ACD41867
45
           901
                    33.4
```

ALIGNMENTS

```
RESULT 1
ABK37922
ID ABK37922 standard; cDNA; 2752 BP.
XX
AC ABK37922;
XX
DT 21-MAY-2002 (first entry)
XX
DE cDNA encoding Human protein NOV1.
XX
KW Human; NOVX; ss; gene; cardiomyopathy; atherosclerosis; diabetes;
```

```
cell signal processing disorder; metabolic disorder; obesity; infection;
KW
     anorexia; cancer-associated cachexia; cancer; neurodegenerative disorder;
KW
     Alzheimer's disease; Parkinson's disease; immune disorder;
KW
     haematopoietic disorders; dyslipidaemia; pain; asthma; hypertension;
KW
     osteoporosis; Crohn's disease; multiple sclerosis; angina pectoris;
KW
     myocardial infarction; ulcer; allergy; benign prostatic hypertrophy;
KW
     psychosis; neurological disorder; anxiety; schizophrenia;
KW
     manic depression; dementia; dyskinesia; Huntington's disease;
KW
     Gilles de la Tourette's syndrome; gene therapy.
KW
XX
os
     Homo sapiens.
XX
PN
     WO200210216-A2.
XX
     07-FEB-2002.
PD
XX
PF
     30-JUL-2001; 2001WO-US024225.
XX
PR
     28-JUL-2000; 2000US-0221409P.
     04-AUG-2000; 2000US-0222840P.
PR
     04-AUG-2000; 2000US-0223752P.
PR
     04-AUG-2000; 2000US-0223762P.
PR
     04-AUG-2000; 2000US-0223769P.
PR
     04-AUG-2000; 2000US-0223770P.
PR
     14-AUG-2000; 2000US-0225146P.
PR
PR
     15-AUG-2000; 2000US-0225392P.
     15-AUG-2000; 2000US-0225470P.
PR
PR
     16-AUG-2000; 2000US-0225697P.
     01-FEB-2001; 2001US-0263662P.
PR
     05-APR-2001; 2001US-0281645P.
PR
XX
PA
     (CURA-) CURAGEN CORP.
XX
     Padigaru M, Mezes P, Mishra V, Burgess C, Casman S, Grosse WM;
PΙ
PΙ
     Alsobrook JP, Lepley DM, Gerlach VL, Macdougall JR, Smithson G;
XX
DR
     WPI; 2002-180074/23.
     P-PSDB; AAU85403.
DR
XX
PT
     New isolated cytoplasmic, nuclear, membrane bound, or secreted
PT
     polypeptide, useful for treating cardiomyopathy, atherosclerosis,
     infections, cancer, neurodegenerative, metabolic, hematopoietic and
PT
PT
     immune disorders.
XX
PS
     Claim 9; Page 9-10; 213pp; English.
XX
     The invention relates to an isolated cytoplasmic, nuclear, membrane
CC
     bound, or secreted polypeptide (NOVX, x= 1-14) their variants or mature
CC
     form. Also included are the nucleic acids encoding the NOVX proteins, a
CC
CC
     vector comprising the nucleic acid, a cell comprising the vector, an anti
CC
     -NOVX antibody and modulators of NOVX. NOVX, the nucleic acid and the
CC
     antibody are useful for treating or preventing a NOVX-associated
CC
     disorder, where the disorder is selected from cardiomyopathy,
CC
     atherosclerosis, diabetes, a disorder related to cell signal processing
CC
     and metabolic pathway modulation, metabolic disorders, obesity,
CC
     infectious disease, anorexia, cancer-associated cachexia, cancer,
CC
     neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,
```

CC immune disorders, haematopoietic disorders, and the various CC dyslipidaemias, metabolic disturbances associated with obesity, the CC metabolic syndrome X and wasting disorders associated with chronic diseases, bacterial, fungal, protozoal and viral infections, pain, CC CC bulimia, asthma, hypertension, urinary retention, osteoporosis, Crohn's CC disease, multiple sclerosis, Albright Hereditary Osteodystrophy, angina CC pectoris, myocardial infarction, ulcer, allergy, benign prostatic CC hypertrophy, and psychotic and neurological disorders, including anxiety, CC schizophrenia, manic depression, delirium, dementia, and dyskinesias, CC such as Huntington's disease and Gilles de la Tourette's syndrome. The nucleic acid is useful in gene therapy. The present sequence encodes a CC CC NOVX protein XX

Sequence 2752 BP; 505 A; 937 C; 829 G; 481 T; 0 U; 0 Other;

SQ

Query Match 100.0%; Score 2697; DB 6; Length 2752;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qу	1	ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC	60
Db	46	ATGGCCGTCCGGCCCGCCTGTGGCCAGCGCTCCTGGCCATAGTCCTCGCCGCTTGGCTC	105
Qy	61	CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG	120
Db	106	CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG	165
Qу	121	GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA	180
Db	166		225
Qy	181	GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG	240
Db	226	GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG	285
Qу	241	TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC	300
Db	286	TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC	345
Qy	301	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG	360
Db	346	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG	405
Qy	361	GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	420
Db	406	GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	465
Qy	421	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
Db	466	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	525
Qy	481	TCCCTGGAGCAGGGCATCGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Db	526	TCCCTGGAGCAGGGCATCGTGCCGTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	585
Qу	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	600

	Db	586	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	645
	Qу	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
	Db	646	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	705
	QУ	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC	720
	Db	706	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTAC	765
٠	Qу	721	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCC	780
	Db	766	GTGAACGGTGGTCGACCGACGTGGACCGACTGTCGCACCCAGCTGTGGGCCC	825
	QУ	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
	Db	826	GGCTGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	885
	Qу	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC	900
	Db	886	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC	945
	Qу	901	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	960
,	Db	946	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	1005
	Qу	961	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGGAGTGCCAGGGCACTGACCTGGAC	1020
	Db	1006	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGGAGTGCCAGGGCACTGACCTGGAC	1065
	Qу	1021	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC	1080
	Db	1066	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC	1125
	Qy .	1081	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATC	1140
	Db	1126	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTCATC	1185
	Qy	1141	CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
	Db	1186	CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1245
	Qу	1201	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
	Db	1246	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1305
	Qy	1261	ACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG	1320
	Db	1306	ACCATCCAGCCGGACCTCAGCACCACCACCACCAGGGGCAGTCTCTGTCCCCGG	1365
	Qу	1321	CAGGATGGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1380
	Db	1366	CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1425
	Qу	1381	GGCGGCCGCCACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC	1440
	Db	1426	GGCGGCCGCCACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC	1485

	•	
	Qy	1441 CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT 1500
	Db	1486 CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT 1545
	Qy	1501 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC 1560
	Db	1546 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC 1605
	Qу	1561 ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1620
	Db	1606 ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1665
	Qу	1621 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1680
	ĎЬ	1666 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1725
	Qy	1681 TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1740
	Db	1726 TGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1785
	Qy	1741 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1800
	Db	1786 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1845
	Qу	1801 GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1860
	Db	1846 GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1905
	QУ	1861 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 1920
	Db	1906 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 1965
	Qy .	1921 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 1980
·	Db	1966 CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 2025
	Qу	1981 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2040
	Db	2026 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2085
	Qу	2041 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2100
	Db	2086 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2145
	Qу	2101 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2160
	Db	2146 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2205
	Qу	2161 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2220
	Db	2206 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2265
	Qy	2221 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2280
	Db	2266 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2325

•

```
2281 GCCTGCAAGCTGTGGGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340
Qу
          2326 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2385
Db
      2341 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400
Qy
          2386 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2445
Db
      2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460
Qу
          2446 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2505
Db
      2461 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
Qy
          2506 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2565
Db
      2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580
Qy
          2566 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2625
Db
      Qу
          Db
      2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
Qу
          Db
      2686 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2742
RESULT 2
ADH71617
   ADH71617 standard; DNA; 2752 BP.
ID
XX
AC
   ADH71617;
XX
   25-MAR-2004 (first entry)
DT
XX
DE
   Human gene of the invention NOV21e SEQ ID NO:513.
XX
KW
   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
   anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW
   vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
KW
   obesity; diabetes; infectious disease; metabolic syndrome X;
KW
   dyslipidaemia.
XX
OS
   Homo sapiens.
XX
PN
   WO2003102155-A2.
XX
PD
   11-DEC-2003.
XX
   03-JUN-2003; 2003WO-US017430.
PF
XX
   03-JUN-2002; 2002US-0385120P.
PR
   04-JUN-2002; 2002US-0385784P.
PR
PR
   05-JUN-2002; 2002US-0386041P.
PR
   05-JUN-2002; 2002US-0386047P.
```

```
06-JUN-2002; 2002US-0386376P.
 PR
      06-JUN-2002; 2002US-0386453P.
 PR
      06-JUN-2002; 2002US-0386864P.
 PR
      06-JUN-2002; 2002US-0387016P.
 PR
      07-JUN-2002; 2002US-0386796P.
 PR
      07-JUN-2002; 2002US-0386816P.
 PR
      07-JUN-2002; 2002US-0386931P.
 PR
      07-JUN-2002; 2002US-0386942P.
 PR
      07-JUN-2002; 2002US-0386971P.
 PR
 PR
      07-JUN-2002; 2002US-0387262P.
      08-JUN-2002; 2002US-0296960P.
· PR
      10-JUN-2002; 2002US-0387400P.
 PR
      10-JUN-2002; 2002US-0387535P.
 PR
      11-JUN-2002; 2002US-0387610P.
 PR
      11-JUN-2002; 2002US-0387625P.
 PR
      11-JUN-2002; 2002US-0387634P.
 PR
      11-JUN-2002; 2002US-0387668P.
 PR
      11-JUN-2002; 2002US-0387696P.
 PR
      11-JUN-2002; 2002US-0387702P.
 PR
      11-JUN-2002; 2002US-0387836P.
 PR
      11-JUN-2002; 2002US-0387859P.
 PR
      12-JUN-2002; 2002US-0387933P.
 PR
      12-JUN-2002; 2002US-0387934P.
 PR
      12-JUN-2002; 2002US-0387960P.
 PR
      12-JUN-2002; 2002US-0388022P.
 PR
      12-JUN-2002; 2002US-0388096P.
 PR
 PR
      13-JUN-2002; 2002US-0389123P.
 PR
      14-JUN-2002; 2002US-0389118P.
      14-JUN-2002; 2002US-0389120P.
 PR
      14-JUN-2002; 2002US-0389144P.
 PR
      14-JUN-2002; 2002US-0389146P.
 PR
      17-JUN-2002; 2002US-0389729P.
 PR
      17-JUN-2002; 2002US-0389742P.
 PR
      18-JUN-2002; 2002US-0389884P.
 PR
      19-JUN-2002; 2002US-0390006P.
 PR
      19-JUN-2002; 2002US-0390209P.
 PR
      21-JUN-2002; 2002US-0390763P.
 PR
      17-JUL-2002; 2002US-0396706P.
 PR
      06-AUG-2002; 2002US-0401628P.
 PR
 PR
      09-AUG-2002; 2002US-0402156P.
      09-AUG-2002; 2002US-0402256P.
 PR
      09-AUG-2002; 2002US-0402389P.
 PR
      12-AUG-2002; 2002US-0402786P.
 PR
      12-AUG-2002; 2002US-0402816P.
 PR
      12-AUG-2002; 2002US-0402821P.
 PR
      12-AUG-2002; 2002US-0402832P.
 PR
      13-AUG-2002; 2002US-0403448P.
 PR
 PR
      13-AUG-2002; 2002US-0403459P.
      13-AUG-2002; 2002US-0403531P.
 PR
      13-AUG-2002; 2002US-0403532P.
 PR
      13-AUG-2002; 2002US-0403563P.
 PR
      13-AUG-2002; 2002US-0406317P.
 PR
      15-AUG-2002; 2002US-0403617P.
 PR
 PR
      26-AUG-2002; 2002US-0406182P.
 PR
      26-AUG-2002; 2002US-0406355P.
 PR
      27-AUG-2002; 2002US-0406240P.
      12-SEP-2002; 2002US-0410084P.
 PR
```

```
20-SEP-2002; 2002US-0412528P.
PR
     23-SEP-2002; 2002US-0412731P.
PR
     30-SEP-2002; 2002US-0414801P.
PR
     30-SEP-2002; 2002US-0414839P.
PR
PR
     30-SEP-2002; 2002US-0414840P.
     30-SEP-2002; 2002US-0414954P.
PR
     09-OCT-2002; 2002US-0417186P.
PR
     09-OCT-2002; 2002US-0417406P.
PR
     23-OCT-2002; 2002US-0420639P.
PR
     28-OCT-2002; 2002US-0421156P.
PR
     31-OCT-2002; 2002US-0422690P.
PR
PR
     01-NOV-2002; 2002US-0423130P.
     05-NOV-2002; 2002US-00423798.
PR
     05-NOV-2002; 2002US-0423798P.
PR
     12-NOV-2002; 2002US-0425453P.
PR
XX
PA
     (CURA-) CURAGEN CORP.
XX
    Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PΙ
     Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI
     Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PΙ
     Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PΙ
    Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PΙ
PΙ
     Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D,
                                                                 Rastelli L;
PΙ
     Rieger DK, Rothenberg ME,
                                 Sciore P, Shenoy SG,
                                                        Shimkets RA;
PΙ
     Smithson G,
                  Spytek KA, Stone DJ, Vernet CAM, Voss EZ,
PΙ
     Zhong H;
XX
DR
     WPI; 2004-081935/08.
DR
     P-PSDB; ADH71618.
XX
    New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT
     treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT
PT
     obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
     Example 21; SEQ ID NO 513; 1880pp; English.
PS
XX
CC
     The invention relates to a novel isolated polypeptide (NOVX). A
CC
     polypeptide of the invention has cytostatic, immunomodulator,
CC
     neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC
     antilipaemic activity, and may have a use in gene therapy, and as a
CC
     vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC
     any of the 303 fully defined nucleotide sequences given in the
CC
     specification. The polypeptide is useful in the manufacture of a
CC
     medicament for treating a syndrome associated with a human disease. The
CC
     polypeptide, polynucleotide and antibody are useful in diagnosing,
CC
     treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC
     Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC
     diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC
     further used as hybridisation probes, in chromosome mapping, tissue
     typing, preventive medicine, and pharmacogenomics. The present sequence
CC
CC
     encodes a NOVX polypeptide of the invention.
XX
SQ
     Sequence 2752 BP; 505 A; 937 C; 829 G; 481 T; 0 U; 0 Other;
  Query Match
                          100.0%;
                                   Score 2697; DB 12; Length 2752;
  Best Local Similarity
                          100.0%; Pred. No. 0;
```

Matches	2697	; Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	ATGGCCGTCCGGCCCGG							60
Db	46	ATGGCCGTCCGGCCCGG							105
Qу	61	CGCGGCTCGGGTGCCCA							120
Db	106	CGCGGCTCGGGTGCCCA							165
Qу	121	GACCTGCTTCCCCACTT							180
Db	166	GACCTGCTTCCCCACTT							225
Qy	181	GTGCTGCTTGTGTGCAA							240
Db	226	GTGCTGCTTGTGTGCAA							285
Qу	241	TGGGTGCGCCAGGTGGA							300
Db	286	TGGGTGCGCCAGGTGG							345
Qy	301	ACCATGGAGGTCCGCAT							360
Db	346	ACCATGGAGGTCCGCAT							405
Qу	361	GAATACTGGTGCCAGTG							420
Db .	406	GAATACTGGTGCCAGTG							465
Qy	421	TACATCCGCATAGCCAC							480
Db	466	TACATCCGCATAGCCA							525
Qу	481	TCCCTGGAGCAGGCAT							540
Db	526	TCCCTGGAGCAGGGCAT							585
Qy	541	GTGGAGTGGCTCCGGAA							600
Db	586	GTGGAGTGGCTCCGGAA	ACGAGG	ACCTGGTGGACCCG	rccc	rggacccca	ATGTA:	TACATC	645
Qу	601	ACGCGGGAGCACAGCCT		- -					660
Db	646	ACGCGGGAGCACAGCCT	rggtgg'	IGCGACAGGCCCGC	CTTG	CTGACACGG	CCAAC'	FACACC	705
Qу	661	TGCGTGGCCAAGAACAT							720
Db	706	TGCGTGGCCAAGAACAT	CGTGG	CACGTCGCCGCAGC	GCCT	CCGCTGCTG'	CATC	GTCTAC	765
Qу	721	GTGAACGGTGGGTGGTC							780
Db	766	GTGAACGGTGGGTGGTC	CGACGT	GGACCGAGTGGTCC	GTCT	GCAGCGCCA	GCTGT	GGGCGC	825
Qу	781	GGCTGGCAGAAACGGAC							840
Db	826	GGCTGGCAGAAACGGAC							885

Qy	841	TGTGAGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC	900
Db	886	${\tt TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC}$	945
Qy	901	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	960
Db	946	${\tt TGGAGCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT}$	1005
Qу	961	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGAC	1020
Db	1006	GAGTGCTCTGACCCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGAC	1065
Qу	1021	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC	1080
Db	1066	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC	1125
Qу	1081	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATC	1140
Db	1126	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGCTGCTGCTGCTTGTCCTCATC	1185
Qу		CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	
Db	1186	CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1245
Qу	1201	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
Db	1246	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1305
Qу	1261	ACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG	1320
Db	1306	ACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG	1365
Qу		CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	
Db	1366	CAGGATGGGCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1425
Qу		GGCGGCCGCCACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC	
Db	1426	GGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC	1485
Qу	1441	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1500
Db	1486	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1545
Qу	1501	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1560
Db	1546	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1605
Qy	1561	ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Db	1606	ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1665
Qу	1621	CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC	1680
Db	1666	CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC	1725

•

	Qy	1681	TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1740
	Db	1726	TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1785
	Qy	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
	Db	1786		1845
	Qу	1801	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG	1860
	Db	1846		1905
•	Qy	1861	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	1920
	Db	1906		1965
	Qy	1921	$\tt CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC$	1980
	Db	1966		2025
	Qy	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2040
	Db	2026	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2085
	Qy	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC	2100
	Db	2086	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC	2145
	Qy	2101	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG	2160
	Db	2146	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG	2205
	Qу	2161	AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG	2220
	Db	2206	AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG	2265
	Qу	2221	CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG	2280
	Db	2266	CAGCGGTACTTGCACCTCCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG	2325
	Qy	2281	GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2340
	Db	2326	GCCTGCAAGCTGTGGGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC	2385
	Qу	2341	AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC	2400
	Db	2386	AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC	2445
	Qу	2401	CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT	2460
	Db	2446	CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT	2505
	Qу	2461	TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC	2520
	Db	2506	TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC	2565
	Qv	2521	CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC	2580

```
2566 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2625
Db
       Qy
           Db
       2641 GCTGGACTGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
Qу
           Db
       2686 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2742
RESULT 3
ABK52891
    ABK52891 standard; DNA; 2697 BP.
ID
XX
AC
   ABK52891;
XX
DΤ
    27-AUG-2002 (first entry)
XX
DE
    Human netrin binding membrane receptor UNC5H-1 DNA sequence #1.
XX
KW
    Netrin binding membrane receptor; receptor; UNC5H-1; gene; ds; human;
    nootropic; neuroprotective; cytostatic; antiparkinsonian;
KW
KW
    cerebroprotective; cancer; central nervous system; CNS; stroke;
    Parkinson's disease; multiple sclerosis; Alzheimer's disease.
KW
XX
os
    Homo sapiens.
XX
FH
    Key
                 Location/Qualifiers
FT
    CDS
                 1. .2697
FΤ
                 /*tag= a
FT
                 /product= "Netrin binding membrane receptor UNC5H-1"
XX
PN
    WO200233080-A2.
XX
PD
    25-APR-2002.
XX
PF
    15-OCT-2001; 2001WO-EP011891.
XX
PR
    16-OCT-2000; 2000US-0240061P.
XX
PA
    (FARB ) BAYER AG.
XX
PΙ
    Koehler RH;
XX
DR
    WPI; 2002-463314/49.
    P-PSDB; AAU97899.
DR
XX
   Novel human netrin binding membrane receptor polypeptide and
PT
    polynucleotides for identifying modulating agents useful in treating
PT
    diseases e.g. Parkinson's disease, multiple sclerosis, stroke,
PT
    Alzheimer's disease.
PT
XX
PS
    Claim 1; Fig 1; 94pp; English.
XX
CC
    This invention relates to the DNA and protein sequences of a novel
```

purified human netrin binding membrane receptor, UNC5H-1. The DNA sequence of the invention is useful as a probe for detecting a nucleic acid encoding the UNC5H-1 protein in a biological sample. The sequences of the invention are useful to screen for agents which decrease the activity of the UNC5H-1 protein. The sequences are also useful for screening agents which regulate (modulate) the activity of the protein of the invention. A pharmaceutical composition containing the protein of the invention or a reagent that modulates the activity of the UNC5H-1 protein may be useful for treating a UNC5H-1 dysfunction related disease such as cancer or a central nervous system (CNS) disorders (e.g, Parkinson's disease, multiple sclerosis, stroke and Alzheimer's disease). Fusion proteins comprising the UNC5H-1 protein are useful for generating antibodies and for in various assay systems, and the protein can be used as a bait protein in a two-hybrid assay or three-hybrid assay. The method of the invention is useful for detecting a coding sequence for the UNC5H-1 protein. The present sequence represents a DNA sequence encoding the human netrin binding membrane receptor UNC5H-1 protein of the invention

SQ Sequence 2697 BP; 503 A; 906 C; 807 G; 481 T; 0 U; 0 Other;

CC

CC XX

> Query Match 99.6%; Score 2687.4; DB 6; Length 2697; Best Local Similarity 99.8%; Pred. No. 0; Matches 2691; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

		.,	
Qy	1	ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC	60
Db	1	ATGGCCGTCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC	60
Qy	61	CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCG	120
Db	61	CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG	120
Qу	121	GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA	180
Db	121	GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA	180
Qy	181	GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG	240
Db	181	GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG	240
Qу	241	TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC	300
Db	241	TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC	300
Qу	301	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG	360
Db	301	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG	360
Qу	361	GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	420
Db	361		420
Qу	421	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
Db	421	TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
Qу	481	TCCCTGGAGCAGGCATCGTGCCCTGCCGTCCACCGGAGGCCATCCCTCCAGCCGAG	540

Db	401		540
Qу		GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	
Db	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	600
Qу	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
Db	601	ACGCGGGACACACCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
Qу	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTAC	720
Db	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTAC	720
Qу	721	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCC	780
Db	721		780
Qу	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
Db	781		840
Qу		TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC	900
Db		TGTGAGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTGGACGGCAGC	900
Qу	901	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	960
Db	901	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	960
Qу	961	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC	1020
Db	961	GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGAC	1020
Qу	1021	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC	1080
Db	1021	ACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTGGCC	1080
Qу	1081	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTGTCCTCATC	1140
Db	1081	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGCTGCTGCTGCTGCTCATC	1140
Qу	1141	$\tt CTCGTTTATTGCCGGAAGAAGGAGGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC$	1200
Db	1141	CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
Qу	1201	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
Db	1201	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
Qy	1261	ACCATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCCCGG	1320
Db	1261	ACCATCCAGCCGGACCTCAGCACCACCACCACCAGGGCAGTCTCTGTCCCCGG	1320
Qу	1321	CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	
		- 3	

Db	1321	CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT	1380
Qy	1381	GGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC	1440
Db	1381	GGCGGCCGCCACACACTGCACCACCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC	1440
Qу	1441	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1500
Db	1441	CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT	1500
Qу	1501	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC	1560
Db	1501	GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTC	1.560
Qу	1561	ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Db	1561	ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG	1620
Qу	1621	CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC	1680
Db	1621	CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC	1680
Qy	1681	TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1740
Db	1681	TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG	1740
Qу	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Db	1741	GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG	1800
Qу	1801	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG	1860
Db	1801	GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG	1860
Qу	1861	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	1920
Db	1861	GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC	1920
Qу	1921	CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	1980
Db	1921	CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC	1980
Qу	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2040
Db	1981	TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG	2040
Qу	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC	2100
Db	2041	GTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC	2100
Qу	2101	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG	2160
Db	2101	TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG	2160
Qу	2161	AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG	2220
Db	2161	AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG	2220

```
2221 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2280
Qy
          2221 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2280
Db
      2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340
Qy
          2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGCCAGAGCTTCAGCATCAACTTC 2340
Db
      2341 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400
Qу
          2341 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400
Db
      2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460
Qу
          2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460
Db
      2461 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
Qy
          2461 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
Db
      2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCACGCCCACAGCCATGATCCTC 2580
Qy
          2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTC 2580
Db
      Qу
          Db
      2641 GCTGGACTGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
Qy
          2641 GCTGGACTGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
Db
RESULT 4
ABK49422
   ABK49422 standard; DNA; 2881 BP.
XX
AC
   ABK49422;
XX
DT
   15-JUL-2002
            (first entry)
XX
   DNA encoding human UNC5-like protein NOV1.
DE
XX
KW
   Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;
   cell signal processing; metabolic pathway modulation; cancerous tissue;
KW
KW
   antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;
KW
   chromosome 13; gene; ds.
XX
os
   Homo sapiens.
XX
               Location/Qualifiers
FH
   Key
               87. .2786
FT
   CDS
FT
               /*tag= a
FT
               /product= "Human UNC5-like protein NOV1"
XX
PN
   WO200229038-A2.
```

```
XX
PD
    11-APR-2002.
XX
    04-OCT-2001; 2001WO-US031377.
PF
XX
    04-OCT-2000; 2000US-0237862P.
PR
XX
PΑ
    (CURA-) CURAGEN CORP.
XX
PΙ
    Herrmann JL, Rastelli L, Shimkets RA;
XX
    WPI; 2002-340104/37.
DR
DR
    P-PSDB; AAU79939.
XX
    Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for
PT
PT
    treating cardiomyopathy, artherosclerosis, and cancer.
XX
PS
    Claim 8; Page 7-8; 180pp; English.
XX
CC
    The present invention relates to a new NOVX polypeptide having a 900
    (NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 865 (NOV5), or 331 (NOV6)
CC
CC
    residue amino acid sequence, as given in the specification. The novel
    polypeptide, and its encoding polynucleotide, are used to treat
CC
    cardiomyopathy, atherosclerosis, cancer or a disease related to cell
CC
CC
    signal processing and metabolic pathway modulation, in a human. Detecting
    the polypeptide or polynucleotide is useful for identifying cancerous
CC
CC
    tissue. The antibody can be used to treat diabetes or cancer. The host
CC
    cells can be used to produce non-human transgenic animals useful in drug
CC
    screening. The present nucleic acid sequence is that of the human UNC5-
    like NOV1 gene located on chromosome 13. This sequence encodes the human
CC
CC
    UNC5-like protein NOV1 of the invention
XX
SQ
    Sequence 2881 BP; 526 A; 985 C; 868 G; 502 T; 0 U; 0 Other;
 Query Match
                       97.28;
                              Score 2621.4; DB 6; Length 2881;
 Best Local Similarity
                       98.9%; Pred. No. 0;
 Matches 2673; Conservative
                             0; Mismatches
                                            21;
                                                Indels
                                                         9;
                                                             Gaps 3;
          1 ATGGCCGTCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
Qу
            87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146
Db
         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
Qу
            147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206
Db
        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
Qу
            207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266
Db
         181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
Qу
            267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326
Db
        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
Qу
            Db
         327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCAGACGGGAGCAGTGGTGAGCCG 386
```

QУ	301	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG	360
Db	387	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG	446
Qy .	361	GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	420
Db	447	GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	506
Qy	421	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
Db	507	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	566
Qу	481	TCCCTGGAGCAGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Db	567	TCCCTGGAGCAGGCATCGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	626
Qу	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	600
Db	627	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATC	686
Qy	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
Db	687	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	746
Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC	720
Db	747	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTCATCGTCTAC	806
Qy	•	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	
Db	807	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	866
Qу		GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	
Db		GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	
Qу		TGTGAGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC	
Db		TGTGAGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC	
Qy .		AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	
Db		AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	
Qy		CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG	
		CGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTG	
~		GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	
		GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	
		GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC	
Db	1167	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTTGTCCTC	1226

Qy	1138	ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1197
Db	1227	ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1286
Qу	1198	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1257
Db	1287	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCAGCAAAGCAGACAACCCCCATCTG	1346
Qу	1258	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCAGGGCAGTCTCTGTCCC	1317
Db	1347	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCC	1403
Qу	1318	CGGCAGGATGGGCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1377
Db	1404	CGGCAGGATGGGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1463
Qу	1378	GGTGGCGCCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1437
Db	1464	GGTGGCGGCCACACACTGCACCACACTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1523
Qу	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1497
Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1583
Qу	1498	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1557
Db	1584	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1643
Qу	1558	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1617
Db	1644	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1703
Qy	1618	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1677
Db	1704		1763
Qy .	1678	AGCTGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1737
Db	1764	AGCTGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1823
Qу	1738	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1797
Db	1824	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1883
Qу	1798	TGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1854
Db	1884	TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1943
Qу	1855	CTGGAGGCCAGTGCCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	1914
Db	1944	CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	2003
Qy	1915	GAGGCCCTCAGCGTGCCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	1974
Db	2004	GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	2063
Qу	1975	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2034

```
Db
     2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2094
Qу
       2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183
Db
     Qy
        Db
     2155 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2214
Qу
        2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303
Db
     2215 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274
Qy
        2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363
Db
     2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334
Qу
        2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423
Db
     2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
Qу
       2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
Db
     2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454
Qy
        2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543
Db
     2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
Qу
        2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603
Db
     2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCACGCCCACAGCCATG 2574
Qу
        2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATG 2663
Db
     Qy
       Db
     2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
Qy
        2724 GCAGTGGCTGGACTGGCCAGCCAGACGCTGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
Db
     2695 TGA 2697
Qу
        111
Db
     2784 TGA 2786
```

RESULT 5 ADH71609

ID ADH71609 standard; DNA; 2881 BP.

XX

AC ADH71609;

```
XX
DT
     25-MAR-2004
                 (first entry)
XX
DE
     Human gene of the invention NOV21a SEQ ID NO:505.
XX
KW
     ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
     anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW
     vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
KW
     obesity; diabetes; infectious disease; metabolic syndrome X;
KW
     dyslipidaemia.
XX
os
     Homo sapiens.
XX
PN
     WO2003102155-A2.
XX
PD
     11-DEC-2003.
XX
PF
     03-JUN-2003; 2003WO-US017430.
XX
PR
     03-JUN-2002; 2002US-0385120P.
PR
     04-JUN-2002; 2002US-0385784P.
     05-JUN-2002; 2002US-0386041P.
PR
     05-JUN-2002; 2002US-0386047P.
PR
     06-JUN-2002; 2002US-0386376P.
PR
PR
     06-JUN-2002; 2002US-0386453P.
     06-JUN-2002; 2002US-0386864P.
PR
PR
     06-JUN-2002; 2002US-0387016P.
PR
     07-JUN-2002; 2002US-0386796P.
     07-JUN-2002; 2002US-0386816P.
PR
     07-JUN-2002; 2002US-0386931P.
PR
     07-JUN-2002; 2002US-0386942P.
PR
     07-JUN-2002; 2002US-0386971P.
PR
     07-JUN-2002; 2002US-0387262P.
PR
PR
     08-JUN-2002; 2002US-0296960P.
PR
     10-JUN-2002; 2002US-0387400P.
     10-JUN-2002; 2002US-0387535P.
PR
     11-JUN-2002; 2002US-0387610P.
PR
     11-JUN-2002; 2002US-0387625P.
PR
     11-JUN-2002; 2002US-0387634P.
PR
     11-JUN-2002; 2002US-0387668P.
PR
     11-JUN-2002; 2002US-0387696P.
PR
PR
     11-JUN-2002; 2002US-0387702P.
PR
     11-JUN-2002; 2002US-0387836P.
     11-JUN-2002; 2002US-0387859P.
PR
     12-JUN-2002; 2002US-0387933P.
PR
     12-JUN-2002; 2002US-0387934P.
PR
     12-JUN-2002; 2002US-0387960P.
PR
PR
     12-JUN-2002; 2002US-0388022P.
PR
     12-JUN-2002; 2002US-0388096P.
PR
     13-JUN-2002; 2002US-0389123P.
PR
     14-JUN-2002; 2002US-0389118P.
     14-JUN-2002; 2002US-0389120P.
PR
     14-JUN-2002; 2002US-0389144P.
PR
     14-JUN-2002; 2002US-0389146P.
PR
PR
     17-JUN-2002; 2002US-0389729P.
PR
     17-JUN-2002; 2002US-0389742P.
PR
     18-JUN-2002; 2002US-0389884P.
```

```
19-JUN-2002; 2002US-0390006P.
PR
PR
     19-JUN-2002; 2002US-0390209P.
PR
     21-JUN-2002; 2002US-0390763P.
     17-JUL-2002; 2002US-0396706P.
PR
PR
     06-AUG-2002; 2002US-0401628P.
     09-AUG-2002; 2002US-0402156P.
PR
     09-AUG-2002; 2002US-0402256P.
PR
     09-AUG-2002; 2002US-0402389P.
PR
     12-AUG-2002; 2002US-0402786P.
PR
     12-AUG-2002; 2002US-0402816P.
PR
     12-AUG-2002; 2002US-0402821P.
PR
PR
     12-AUG-2002; 2002US-0402832P.
PR
     13-AUG-2002; 2002US-0403448P.
     13-AUG-2002; 2002US-0403459P.
PR
     13-AUG-2002; 2002US-0403531P.
PR
     13-AUG-2002; 2002US-0403532P.
PR
     13-AUG-2002; 2002US-0403563P.
PR
     13-AUG-2002; 2002US-0406317P.
PR
     15-AUG-2002; 2002US-0403617P.
PR
     26-AUG-2002; 2002US-0406182P.
PR
     26-AUG-2002; 2002US-0406355P.
PR
     27-AUG-2002; 2002US-0406240P.
PR
     12-SEP-2002; 2002US-0410084P.
PR
PR
     20-SEP-2002; 2002US-0412528P.
PR
     23-SEP-2002; 2002US-0412731P.
PR
     30-SEP-2002; 2002US-0414801P.
     30-SEP-2002; 2002US-0414839P.
PR
     30-SEP-2002; 2002US-0414840P.
PR
     30-SEP-2002; 2002US-0414954P.
PR
     09-OCT-2002; 2002US-0417186P.
PR
     09-OCT-2002; 2002US-0417406P.
PR
     23-OCT-2002; 2002US-0420639P.
PR
PR
     28-OCT-2002; 2002US-0421156P.
PR
     31-OCT-2002; 2002US-0422690P.
     01-NOV-2002; 2002US-0423130P.
PR
     05-NOV-2002; 2002US-00423798.
PR
     05-NOV-2002; 2002US-0423798P.
PR
PR
     12-NOV-2002; 2002US-0425453P.
XX
PA
     (CURA-) CURAGEN CORP.
XX
PΙ
     Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PΙ
     Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
                   Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI
     Ettenberg S,
     Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PΙ
PΙ
     Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI
     Padigaru M,
                  Patturajan M, Pena CEA, Peyman JA, Raha D,
                                                                 Rastelli L;
PΙ
     Rieger DK, Rothenberg ME,
                                 Sciore P, Shenoy SG, Shimkets RA;
PΙ
     Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ,
PΙ
     Zhong H;
XX
     WPI; 2004-081935/08.
DR
DR
     P-PSDB; ADH71610.
XX
PT
     New NOVX polypeptides and nucleic acid molecules useful for preventing or
```

treating NOVX-associated disorders, e.g. cancer, diabetes, infection or obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

PT

PT

```
XX
PS Example 21; SEQ ID NO 505; 1880pp; English.

XX
CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobic
```

polypeptide of the invention has cytostatic, immunomodulator, neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and antilipaemic activity, and may have a use in gene therapy, and as a vaccine. The polypeptides are encoded by NOVX polynucleotides comprising any of the 303 fully defined nucleotide sequences given in the specification. The polypeptide is useful in the manufacture of a medicament for treating a syndrome associated with a human disease. The polypeptide, polynucleotide and antibody are useful in diagnosing, treating or preventing NOVX-associated disorders, e.g. cancer, cachexia, Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are further used as hybridisation probes, in chromosome mapping, tissue typing, preventive medicine, and pharmacogenomics. The present sequence encodes a NOVX polypeptide of the invention.

CC XX SO

CC

Sequence 2881 BP; 526 A; 985 C; 868 G; 502 T; 0 U; 0 Other;

97.2%; Query Match Score 2621.4; DB 12; Length 2881; 98.9%; Best Local Similarity Pred. No. 0; Matches 2673: Conservative Mismatches 21: Indels 9: Gaps 3; 1 ATGGCCGTCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTTGGCTC 60 Qу

Db 87 ATGGCCGTCCGGCCTGTGGCCAGCGCTCTGGGCATAGTCCTCGCCGCTTGGCTC 146 61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120 Qy 147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206 Db 121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180 Qy 207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266 Db 181 GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240 Qv 267 GTGCTGCTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326 Db 241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCAGACGGGAGCAGTGGGCTGCCC 300 Qv 327 TGGGTGCGCCAGGTGACCACGTGATCGAGCGCAGCAGACGGGAGCAGTGGTGAGCCG 386 Db 301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360 Qу Db Qу

	Qÿ	401		540
	Db	567	TCCCTGGAGCAGGCATCGTGCTGCCCTGCCGTCCACCGGAGGCCATCCCTCCAGCCGAG	626
	Qy	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	600
	Db	627	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	686
	Qy	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
	Db	687		746
	Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC	720
	Db	747	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTAC	806
	Qy	721	GTGAACGGTGGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCC	780
	Db	807	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	866
	Qy	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
*	Db	867	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	926
	Qy	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC	897
	Db	927	TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTGCTTGTCTCTGTGGACGGC	986
	Qy	898	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	957
	Db	987	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	1046
	Qy	958	CGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGGGAGG	1017
	Db	1047	CGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGGGAGG	1106
	Qy	1018	GACACCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1077
	Db	1107	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1166
	Qy	1078	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC	1137
	Db	1167	GCCCTCTATGTGGGCCCTCATCGCCGTGGCCGTCTGCCTGC	1226
•	Qу	1138	ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1197
	Db	1227	ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1286
	Qy	1198	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1257
	Db	1287		1346
	Qy	1258	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCAGGGCAGTCTCTGTCCC	1317
	Db	1347		1403
•	04	1210	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	1277

D.L.	1404		1462
Db	1404	CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1463
Qу	1378	GGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1437
Db	1464	GGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1523
Qу	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1497
Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1583
Qу	1498	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1557
Db	1584	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1643
Qу	1558	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1617
Db	1644	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1703
Qу	1618	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1677
Db	1704	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1763
Qy	1678	AGCTGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1737
Db	1764	AGCTGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1823
Qу	1738	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1797
Db	1824	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1883
Qy	1798	TGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1854
Db	1884	${\tt TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG}$	1943
Qy	1855	CTGGAGGCCAGTGCCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	1914
Db	1944 [.]	CTGGAGGCCAGTGCCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	2003
QУ		GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	
Db	2004	GAGGCCCTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	2063
Qу	1975	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2034
Db	2064	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2123
Qу	2035	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC	2094
Db		AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC	
Qу	2095	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	2154
Db	2184	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	2243
Qу	2155	CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT	2214

```
2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303
Db
      2215 GGCACGCAGCGGTACTTGCACTGCACCTTGACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274
Qv
          2304 GGCACGCAGCGGTACTTGCACTGCACCTTGACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363
Db
      2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334
Qу
          2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGCCGACGGCAGAGCTTCAGCATC 2423
Db
      2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
Qу
          2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
Db
      2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454
Qy
          2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543
Db
      2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
Qy
          2544 ATAATTTCCAGCCTGGACCCACCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603
Db
      2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATG 2574
Qy
          2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATG 2663
Dh
      Qу
          Nh
      2635 GCAGTGGCTGGACTGGCCAGCCAGACGCTGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
Qу
          2724 GCAGTGGCTGGACTGGCCAGCCAGACGCTGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
Db
      2695 TGA 2697
Qу
          \mathbf{I} \mathbf{I} \mathbf{I}
      2784 TGA 2786
Db
RESULT 6
ADH71633
   ADH71633 standard; DNA; 2880 BP.
XX
AC
   ADH71633;
XX
DT
   25-MAR-2004 (first entry)
XX
   Human gene of the invention NOV21m SEQ ID NO:529.
DE
XX
   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW
   anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW
   vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
KW
   obesity; diabetes; infectious disease; metabolic syndrome X;
KW
   dyslipidaemia.
XX
os
   Homo sapiens.
XX
```

```
PN
     WO2003102155-A2.
XX
PD
     11-DEC-2003.
XX
     03-JUN-2003; 2003WO-US017430.
PF
XX
     03-JUN-2002; 2002US-0385120P.
PR
     04-JUN-2002; 2002US-0385784P.
PR
PR
     05-JUN-2002; 2002US-0386041P.
PR
     05-JUN-2002; 2002US-0386047P.
     06-JUN-2002; 2002US-0386376P.
PR
     06-JUN-2002; 2002US-0386453P.
PR
PR
     06-JUN-2002; 2002US-0386864P.
     06-JUN-2002; 2002US-0387016P.
PR
     07-JUN-2002; 2002US-0386796P.
PR
     07-JUN-2002; 2002US-0386816P.
PR
     07-JUN-2002; 2002US-0386931P.
PR
     07-JUN-2002; 2002US-0386942P.
PR
     07-JUN-2002; 2002US-0386971P.
PR
     07-JUN-2002; 2002US-0387262P.
PR
     08-JUN-2002; 2002US-0296960P.
PR
     10-JUN-2002; 2002US-0387400P.
PR
     10-JUN-2002; 2002US-0387535P.
PR
     11-JUN-2002; 2002US-0387610P.
PR
     11-JUN-2002; 2002US-0387625P.
PR
     11-JUN-2002; 2002US-0387634P.
PR
PR
     11-JUN-2002; 2002US-0387668P.
PR
     11-JUN-2002; 2002US-0387696P.
     11-JUN-2002; 2002US-0387702P.
PR
     11-JUN-2002; 2002US-0387836P.
PR
     11-JUN-2002; 2002US-0387859P.
PR
     12-JUN-2002; 2002US-0387933P.
PR
     12-JUN-2002; 2002US-0387934P.
PR
     12-JUN-2002; 2002US-0387960P.
PR
PR
     12-JUN-2002; 2002US-0388022P.
     12-JUN-2002; 2002US-0388096P.
PR
     13-JUN-2002; 2002US-0389123P.
PR
     14-JUN-2002; 2002US-0389118P.
PR
PR
     14-JUN-2002; 2002US-0389120P.
PR
     14-JUN-2002; 2002US-0389144P.
     14-JUN-2002; 2002US-0389146P.
PR
     17-JUN-2002; 2002US-0389729P.
PR
     17-JUN-2002; 2002US-0389742P.
PR
     18-JUN-2002; 2002US-0389884P.
PR
     19-JUN-2002; 2002US-0390006P.
PR
     19-JUN-2002; 2002US-0390209P.
PR
     21-JUN-2002; 2002US-0390763P.
PR
     17-JUL-2002; 2002US-0396706P.
PR
     06-AUG-2002; 2002US-0401628P.
PR
     09-AUG-2002; 2002US-0402156P.
PR
     09-AUG-2002; 2002US-0402256P.
PR
     09-AUG-2002; 2002US-0402389P.
PR
     12-AUG-2002; 2002US-0402786P.
PR
     12-AUG-2002; 2002US-0402816P.
PR
PR
     12-AUG-2002; 2002US-0402821P.
PR
     12-AUG-2002; 2002US-0402832P.
     13-AUG-2002; 2002US-0403448P.
PR
```

```
PR
     13-AUG-2002; 2002US-0403459P.
     13-AUG-2002; 2002US-0403531P.
PR
PR
     13-AUG-2002; 2002US-0403532P.
PR
     13-AUG-2002; 2002US-0403563P.
PR
     13-AUG-2002; 2002US-0406317P.
PR
     15-AUG-2002; 2002US-0403617P.
     26-AUG-2002; 2002US-0406182P.
PR
     26-AUG-2002; 2002US-0406355P.
PR
PR
     27-AUG-2002; 2002US-0406240P.
PR
     12-SEP-2002; 2002US-0410084P.
     20-SEP-2002; 2002US-0412528P.
PR
     23-SEP-2002; 2002US-0412731P.
PR
     30-SEP-2002; 2002US-0414801P.
PR
     30-SEP-2002; 2002US-0414839P.
PR
     30-SEP-2002; 2002US-0414840P.
PR
     30-SEP-2002; 2002US-0414954P.
PR
     09-OCT-2002; 2002US-0417186P.
PR
     09-OCT-2002; 2002US-0417406P.
PR
     23-OCT-2002; 2002US-0420639P.
PR
     28-OCT-2002; 2002US-0421156P.
PR
     31-OCT-2002; 2002US-0422690P.
PR
     01-NOV-2002; 2002US-0423130P.
PR
     05-NOV-2002; 2002US-00423798.
PR
     05-NOV-2002; 2002US-0423798P.
PR
     12-NOV-2002; 2002US-0425453P.
PR
XX
PΑ
     (CURA-) CURAGEN CORP.
XX
PΙ
     Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
     Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PΙ
PΙ
     Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PΙ
     Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
    Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PΙ
PΙ
     Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D,
     Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PΙ
PΙ
     Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
ΡI
     Zhong H;
XX
    WPI; 2004-081935/08.
DR
DR
     P-PSDB; ADH71634.
XX
     New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT
     treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT
     obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
PT ·
XX
PS
     Example 21; SEQ ID NO 529; 1880pp; English.
XX
     The invention relates to a novel isolated polypeptide (NOVX). A
CC
CC
     polypeptide of the invention has cytostatic, immunomodulator,
CC
     neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC
     antilipaemic activity, and may have a use in gene therapy, and as a
CC
     vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC
     any of the 303 fully defined nucleotide sequences given in the
CC
     specification. The polypeptide is useful in the manufacture of a
CC
     medicament for treating a syndrome associated with a human disease. The
CC
     polypeptide, polynucleotide and antibody are useful in diagnosing,
CC
     treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
```

```
Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC
     diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC
CC
     further used as hybridisation probes, in chromosome mapping, tissue
     typing, preventive medicine, and pharmacogenomics. The present sequence
CC
     encodes a NOVX polypeptide of the invention.
CC
XX
     Sequence 2880 BP; 527 A; 984 C; 867 G; 502 T; 0 U; 0 Other;
SO
  Query Match
                          97.1%;
                                  Score 2619.8;
                                                 DB 12; Length 2880;
```

Best Local Similarity 98.9%; Pred. No. 0; Matches 2672; Conservative Mismatches 22: Indels 9; 3; 0; Gaps 1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60 Qу 86 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 145 Db 61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120 Qу 146 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 205 Db 121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180 Qy 206 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 265 Db 181 GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240 Qу 266 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 325 Db 241 TGGGTGCGCCAGGTGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300 Qу 326 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 385 Db 301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360 Qу 386 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 445 Db 361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420 Qу Db 421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480 Qу 506 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 565 Db 481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540 Qу 566 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCCATCCCTCCAGCCGAG 625 Db 541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600 Qу 626 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 685 Db 601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660 Qy 686 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 745 Db 661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720 Qy

Db	746	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTCATCGTCTAC	805
Qy	721	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	780
Db	806	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCC	865
Qу	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
Db	866	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	925
Qу	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC	897
Db	926	TGTGAGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC	985
Qу	898	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	957
Db	986	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	1045
Qу	958	CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG	1017
Db	1046	CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTG	1105
Qу	1018	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1077
Db	1106	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1165
Qу	1078	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC	1137
Db	1166	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC	1225
Qy .		ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1197
Db			1285
Qy	1198	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1257
Db	1286	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCAGCAAAGCAGACAACCCCCATCTG	1345
Qу	1258	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCAGGGCAGTCTCTGTCCC	1317
Db	1346	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCC	1402
Qу	1318	CGGCAGGATGGGCCCCAGCTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1377
Db	1403	CGGCAGGATGGGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1462
Qy	1378	GGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1437
Db	1463	GGTGGCGGCCACACACTGCACCACACTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1522
Qу	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1497
Db	1523	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1582
Qу	1498	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1557

	Db	1583	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1642
	Qу	1558	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1617
	Db	1643	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1702
	Qy	1618	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1677
	Db	1703	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1762
	Qу	1678	AGCTGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1737
	Db	1763	AGCTGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1822
	Qу	1738	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1797
	Db	1823	GGGGAGCCCAGCCTGACAGCTGGAGCCTGCACCTCAAAAAGCAGTCGTGCGAGGGCAGC	1882
	Qy	1798	TGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1854
	Db	1883	TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1942
	Qу	1855	CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	1914
	Db	1943	${\tt CTGGAGGCCAGTGCCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA}$	2002
	Qy	1915	GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	1974
	Db	2003	GAGGCCCTCAGCGTGGCTGCCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	2062
	Qу		TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	
	Db		TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	
	Qу		AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC	
	Db		AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC	
	QУ		CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	
	Db		CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	
•	QУ		CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT	
	Db		CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT	
	ДУ		GGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT	
	Db		GGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT	
	QУ		GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC	
	Db		GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGGGGAGGGCGACGGGCAGAGCTTCAGCATC	
	Qу		AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	
	Db	2423	AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2482

```
Qу
       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454
           2483 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2542
Db
       2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
Qу
           2543 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2602
Db
       2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATG 2574
Qу
           2603 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATG 2662
Db
       Qу
           Db
       2635 GCAGTGGCTGGACTGGCCAGCCAGACGCTGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
Qу
           2723 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2782
Db
       2695 TGA 2697
Qу
           III
Db
       2783 TGA 2785
RESULT 7
ADH71649
ID
   ADH71649 standard; DNA; 2881 BP.
XX
AC
   ADH71649;
XX
DT
   25-MAR-2004 (first entry)
XX
DE
   Human gene of the invention NOV21u SEQ ID NO:545.
XX
KW
   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
    anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW
    vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
KW
    obesity; diabetes; infectious disease; metabolic syndrome X;
KW
   dyslipidaemia.
XX
os
   Homo sapiens.
XX
PN
   WO2003102155-A2.
XX
PD
   11-DEC-2003.
XX
    03-JUN-2003; 2003WO-US017430.
PF
XX
    03-JUN-2002; 2002US-0385120P.
PR
    04-JUN-2002; 2002US-0385784P.
PR
    05-JUN-2002; 2002US-0386041P.
PR
    05-JUN-2002; 2002US-0386047P.
PR
    06-JUN-2002; 2002US-0386376P.
PR
    06-JUN-2002; 2002US-0386453P.
PR
    06-JUN-2002; 2002US-0386864P.
PR
```

```
06-JUN-2002; 2002US-0387016P.
PR
     07-JUN-2002; 2002US-0386796P.
PR
PR
     07-JUN-2002; 2002US-0386816P.
     07-JUN-2002; 2002US-0386931P.
PR
     07-JUN-2002; 2002US-0386942P.
PR
     07-JUN-2002; 2002US-0386971P.
PR
     07-JUN-2002; 2002US-0387262P.
PR
     08-JUN-2002; 2002US-0296960P.
PR
PR
     10-JUN-2002; 2002US-0387400P.
PR
     10-JUN-2002; 2002US-0387535P.
     11-JUN-2002; 2002US-0387610P.
PR
     11-JUN-2002; 2002US-0387625P.
PR
     11-JUN-2002; 2002US-0387634P.
PR
     11-JUN-2002; 2002US-0387668P.
PR
     11-JUN-2002; 2002US-0387696P.
PR
     11-JUN-2002; 2002US-0387702P.
PR
     11-JUN-2002; 2002US-0387836P.
PR
PR
     11-JUN-2002; 2002US-0387859P.
PR
     12-JUN-2002; 2002US-0387933P.
     12-JUN-2002; 2002US-0387934P.
PR
     12-JUN-2002; 2002US-0387960P.
PR
     12-JUN-2002; 2002US-0388022P.
PR
PR
     12-JUN-2002; 2002US-0388096P.
PR
     13-JUN-2002; 2002US-0389123P.
     14-JUN-2002; 2002US-0389118P.
PR
PR
     14-JUN-2002; 2002US-0389120P.
     14-JUN-2002; 2002US-0389144P.
PR
     14-JUN-2002; 2002US-0389146P.
PR
     17-JUN-2002; 2002US-0389729P.
PR
     17-JUN-2002; 2002US-0389742P.
PR
     18-JUN-2002; 2002US-0389884P.
PR
PR
     19-JUN-2002; 2002US-0390006P.
     19-JUN-2002; 2002US-0390209P.
PR
PR
     21-JUN-2002; 2002US-0390763P.
     17-JUL-2002; 2002US-0396706P.
PR
     06-AUG-2002; 2002US-0401628P.
PR
     09-AUG-2002; 2002US-0402156P.
PR
     09-AUG-2002; 2002US-0402256P.
PR
     09-AUG-2002; 2002US-0402389P.
PR
     12-AUG-2002; 2002US-0402786P.
PR
PR
     12-AUG-2002; 2002US-0402816P.
PR
     12-AUG-2002; 2002US-0402821P.
     12-AUG-2002; 2002US-0402832P.
PR
     13-AUG-2002; 2002US-0403448P.
PR
     13-AUG-2002; 2002US-0403459P.
PR
PR
     13-AUG-2002; 2002US-0403531P.
PR
     13-AUG-2002; 2002US-0403532P.
PR
     13-AUG-2002; 2002US-0403563P.
PR
     13-AUG-2002; 2002US-0406317P.
PR
     15-AUG-2002; 2002US-0403617P.
     26-AUG-2002; 2002US-0406182P.
PR
     26-AUG-2002; 2002US-0406355P.
PR
     27-AUG-2002; 2002US-0406240P.
PR
     12-SEP-2002; 2002US-0410084P.
PR
PR
     20-SEP-2002; 2002US-0412528P.
     23-SEP-2002; 2002US-0412731P.
PR
PR
     30-SEP-2002; 2002US-0414801P.
```

```
30-SEP-2002; 2002US-0414839P.
PR
     30-SEP-2002; 2002US-0414840P.
PR
     30-SEP-2002; 2002US-0414954P.
PR
     09-OCT-2002; 2002US-0417186P.
PR
     09-OCT-2002; 2002US-0417406P.
PR
PR
     23-OCT-2002; 2002US-0420639P.
     28-OCT-2002; 2002US-0421156P.
PR
     31-OCT-2002; 2002US-0422690P.
PR
     01-NOV-2002; 2002US-0423130P.
PR
PR
     05-NOV-2002; 2002US-00423798.
     05-NOV-2002; 2002US-0423798P.
PR
     12-NOV-2002; 2002US-0425453P.
PR
XX
PA
     (CURA-) CURAGEN CORP.
XX
    Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PΙ
    Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PΙ
     Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PΙ
     Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PΙ
    Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PΙ
     Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D,
                                                                Rastelli L;
PΙ
     Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PΙ
     Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PΙ
PΙ
     Zhong H;
XX
DR
    WPI; 2004-081935/08.
DR
     P-PSDB; ADH71650.
XX
PT
    New NOVX polypeptides and nucleic acid molecules useful for preventing or
     treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT
PT
     obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS
     Example 21; SEQ ID NO 545; 1880pp; English.
XX
CC
    The invention relates to a novel isolated polypeptide (NOVX). A
CC
     polypeptide of the invention has cytostatic, immunomodulator,
     neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC
CC
     antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
     any of the 303 fully defined nucleotide sequences given in the
CC
CC
     specification. The polypeptide is useful in the manufacture of a
CC
     medicament for treating a syndrome associated with a human disease. The
CC
     polypeptide, polynucleotide and antibody are useful in diagnosing,
CC
     treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
     Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC
     diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC
CC
     further used as hybridisation probes, in chromosome mapping, tissue
CC
     typing, preventive medicine, and pharmacogenomics. The present sequence
CC
     encodes a NOVX polypeptide of the invention.
XX
     Sequence 2881 BP; 526 A; 986 C; 868 G; 501 T; 0 U; 0 Other;
SQ
  Query Match
                          97.1%;
                                  Score 2619.8; DB 12; Length 2881;
  Best Local Similarity
                          98.9%;
                                  Pred. No. 0;
  Matches 2672; Conservative
                                 0; Mismatches
                                                  22; Indels
                                                                     Gaps
                                                                             3;
```

Db	87	ATGGCCGTCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC	146
Qy	61	CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG	120
Db	147	CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG	206
Qy	121	GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA	180
Db	207	GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA	266
Qy	181	GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG	240
Db	267	GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG	326
Qy	241	TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC	300
Db	327	TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG	386
Qy	301	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG	360
Db	387	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG	446
Qу	361	GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	420
Db	447	GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	506
Qу	421	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
Db	507	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	566
Qy	481	TCCCTGGAGCAGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Db	567	TCCCTGGAGCAGGCATCGTGCCGTGCCGTCCACCGGAGGCCATCCCTCCAGCCGAG	626
Qy	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	600
Db	627	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	686
Qy	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
Db	687	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	746
Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC	720
Db	747	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC	806
Qy	721	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	780
Db	807	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCC	866
Qy	781	GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
Db	867	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	926
Qy	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC	897

Db	927	TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC	986
Qy	898	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	957
Db	987	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	1046
Qy	958	CGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGGGAGG	1017
Db	1047	CGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGGGAGTGCCAGGGCACTGACCTG	1106
Qy	1018	GACACCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1077
Db	1107	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1166
Qy	1078	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC	1137
Db	1167	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC	1226
Qy	1138	ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1197
Db	1227	ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1286
Qу	1198	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1257
Db	1287	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1346
Qy .	1258	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCAGGGCAGTCTCTGTCCC	1317
Db	1347	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCC	1403
Qу	1318	CGGCAGGATGGGCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1377
Db	1404	CGGCAGGATGGGCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1463
Qу	1378	GGTGGCGCCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1437
Db	1464	GGTGGCGGCCACACACTGCACCACACTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1523
Qy	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1497
Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1583
Qy	1498	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1557
Db	1584	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1643
Qy	1558	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1617
Db	1644	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1703
Qу	1618	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1677
Db	1704	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1763
Qу	1678	AGCTGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1737
Db	1764	AGCTGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1823

•	Qу	1738	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1797
	Db	1824		1883
•	Qy	1798	TGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1854
	Db	1884	TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1943
	Qy	1855	CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	1914
:	Db	1944	CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	2003
•	Qy	1915	GAGGCCCTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	1974
:	Db :	2004	GAGGCCTCAGCGTGGCTGCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	2063
•	Qy		TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2034
:	Db :		TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2123
٠ (Qy :	2035	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC	2094
:	Db :	2124	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC	2183
•	Qy :	2095	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	2154
1	Db :	2184	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	2243
•	Qy :	2155	CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT	2214
1	Db :	2244	CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT	2303
(Qy :	2215	GGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT	2274
1	Db :	2304	GGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT	2363
(Qy :	2275	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC	2334
1	Db :	2364	GACCTGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC	2423
(ÖÄ :	2335	AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGGCTCTGGAGAGTGAAGCG	2394
1	Db :	2424	AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2483
(Qy :	2395	GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG	2454
• :	Db :	2484	GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG	2543
•	Qy :	2455	ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG	2514
:	Db :	2544	ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG	2603
•	Qy :	2515	AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATG	2574 .
:	Db :	2604	AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG	2663

```
Qу
       Db
       2635 GCAGTGGCTGGACTGGGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
Qy
            Db
       2724 GCAGTGGCTGGACTGGGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
       2695 TGA 2697
Qy
            \Pi
       2784 TGA 2786
Db
RESULT 8
ADH71635
    ADH71635 standard; DNA; 2881 BP.
XX
AC
    ADH71635;
XX
DT
    25-MAR-2004 (first entry)
XX
DE
    Human gene of the invention NOV21n SEQ ID NO:531.
XX
KW
    ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW
    anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW
    vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
    obesity; diabetes; infectious disease; metabolic syndrome X;
KW
    dyslipidaemia.
XX
os
    Homo sapiens.
XX
    WO2003102155-A2.
PN
XX
    11-DEC-2003.
PD
XX
    03-JUN-2003; 2003WO-US017430.
PF
XX
PR
    03-JUN-2002; 2002US-0385120P.
    04-JUN-2002; 2002US-0385784P.
PR
    05-JUN-2002; 2002US-0386041P.
PR
    05-JUN-2002; 2002US-0386047P.
PR
    06-JUN-2002; 2002US-0386376P.
PR
PR
    06-JUN-2002; 2002US-0386453P.
    06-JUN-2002; 2002US-0386864P.
PR
    06-JUN-2002; 2002US-0387016P.
PR
    07-JUN-2002; 2002US-0386796P.
PR
    07-JUN-2002; 2002US-0386816P.
PR
PR
    07-JUN-2002; 2002US-0386931P.
PR
    07-JUN-2002; 2002US-0386942P.
PR
    07-JUN-2002; 2002US-0386971P.
PR
    07-JUN-2002; 2002US-0387262P.
    08-JUN-2002; 2002US-0296960P.
PR
    10-JUN-2002; 2002US-0387400P.
PR
    10-JUN-2002; 2002US-0387535P.
PR
PR
    11-JUN-2002; 2002US-0387610P.
    11-JUN-2002; 2002US-0387625P.
PR
PR
    11-JUN-2002; 2002US-0387634P.
```

```
11-JUN-2002; 2002US-0387668P.
PR
     11-JUN-2002; 2002US-0387696P.
PR
     11-JUN-2002; 2002US-0387702P.
PR
     11-JUN-2002; 2002US-0387836P.
PR
PR
     11-JUN-2002; 2002US-0387859P.
PR
     12-JUN-2002; 2002US-0387933P.
     12-JUN-2002; 2002US-0387934P.
PR
     12-JUN-2002; 2002US-0387960P.
PR
     12-JUN-2002; 2002US-0388022P.
PR
     12-JUN-2002; 2002US-0388096P.
PR
     13-JUN-2002; 2002US-0389123P.
PR
     14-JUN-2002; 2002US-0389118P.
PR
     14-JUN-2002; 2002US-0389120P.
PR
     14-JUN-2002; 2002US-0389144P.
PR
     14-JUN-2002; 2002US-0389146P.
PR
     17-JUN-2002; 2002US-0389729P.
PR
     17-JUN-2002; 2002US-0389742P.
PR
     18-JUN-2002; 2002US-0389884P.
PR
     19-JUN-2002; 2002US-0390006P.
PR
     19-JUN-2002; 2002US-0390209P.
PR
     21-JUN-2002; 2002US-0390763P.
PR
     17-JUL-2002; 2002US-0396706P.
PR
PR
     06-AUG-2002; 2002US-0401628P.
PR
     09-AUG-2002; 2002US-0402156P.
PR
     09-AUG-2002; 2002US-0402256P.
PR
     09-AUG-2002; 2002US-0402389P.
     12-AUG-2002; 2002US-0402786P.
PR
     12-AUG-2002; 2002US-0402816P.
PR
     12-AUG-2002; 2002US-0402821P.
PR
     12-AUG-2002; 2002US-0402832P.
PR
     13-AUG-2002; 2002US-0403448P.
PR
     13-AUG-2002; 2002US-0403459P.
PR
     13-AUG-2002; 2002US-0403531P.
PR
PR
     13-AUG-2002; 2002US-0403532P.
     13-AUG-2002; 2002US-0403563P.
PR
     13-AUG-2002; 2002US-0406317P.
PR
     15-AUG-2002; 2002US-0403617P.
PR
     26-AUG-2002; 2002US-0406182P.
PR
     26-AUG-2002; 2002US-0406355P.
PR
     27-AUG-2002; 2002US-0406240P.
PR
     12-SEP-2002; 2002US-0410084P.
PR
PR
     20-SEP-2002; 2002US-0412528P.
     23-SEP-2002; 2002US-0412731P.
PR
     30-SEP-2002; 2002US-0414801P.
PR
     30-SEP-2002; 2002US-0414839P.
PR
     30-SEP-2002; 2002US-0414840P.
PR
     30-SEP-2002; 2002US-0414954P.
PR
PR
     09-OCT-2002; 2002US-0417186P.
PR
     09-OCT-2002; 2002US-0417406P.
     23-OCT-2002; 2002US-0420639P.
PR
     28-OCT-2002; 2002US-0421156P.
PR
     31-OCT-2002; 2002US-0422690P.
PR
PR
     01-NOV-2002; 2002US-0423130P.
PR
     05-NOV-2002; 2002US-00423798.
PR
     05-NOV-2002; 2002US-0423798P.
     12-NOV-2002; 2002US-0425453P.
PR
XX
```

```
(CURA-) CURAGEN CORP.
PA
XX
PΙ
    Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
    Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI
PΙ
    Ettenberg S, Gangolli EA,
                              Gerlach VL, Gorman L, Gunther E,
PΙ
    Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
ΡI
    Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
    Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PΙ
PΙ
    Rieger DK, Rothenberg ME,
                              Sciore P, Shenoy SG, Shimkets RA;
PΙ
    Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PΙ
    Zhong H;
XX
DR
    WPI; 2004-081935/08.
DR
    P-PSDB; ADH71636.
XX
PT
    New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT
    treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
    obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
PT
XX
PS
    Example 21; SEQ ID NO 531; 1880pp; English.
XX
CC
    The invention relates to a novel isolated polypeptide (NOVX). A
CC
    polypeptide of the invention has cytostatic, immunomodulator,
CC
    neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC
    antilipaemic activity, and may have a use in gene therapy, and as a
CC
    vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC
    any of the 303 fully defined nucleotide sequences given in the
CC
    specification. The polypeptide is useful in the manufacture of a
CC
    medicament for treating a syndrome associated with a human disease. The
CC
    polypeptide, polynucleotide and antibody are useful in diagnosing,
    treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC
    Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC
CC
    diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC
    further used as hybridisation probes, in chromosome mapping, tissue
CC
    typing, preventive medicine, and pharmacogenomics. The present sequence
CC
    encodes a NOVX polypeptide of the invention.
XX
    Sequence 2881 BP; 527 A; 985 C; 867 G; 502 T; 0 U; 0 Other;
SQ
                        97.1%;
                               Score 2619.8; DB 12; Length 2881;
 Query Match
 Best Local Similarity
                        98.9%;
                               Pred. No. 0;
 Matches 2672; Conservative
                                  Mismatches
                                                                Gaps
                                                                        3;
                                                   Indels
           1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
Qу
             Db
          87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146
          61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
Qу
             147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206
Db
         121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
Qу
             207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266
Db
Qу
         181 GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
```

Db	267	GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG	326
Qy	241	TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC	300
Db	327	TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG	386
Qy	301	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG	.360
Db	387	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG	446
Qy	361	GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	420
Db	447		506
Qy	421	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
Db	507	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	566
Qy	481	TCCCTGGAGCAGGCATCGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Db	567	TCCCTGGAGCAGGCATCGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	626
Qу	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	600
Db	627	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATC	686.
Qy	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
Db	687	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	746
Qy	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC	720
Db	747	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC	806
Qy	721	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	780
Db	807	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCC	866
Qy	781	GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
Db	867	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	926
Qy	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC	897
Db	927	TGTGAGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGTGTCTCTGTGGACGGC	986
Qy	898	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	957
Db	987	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	1046
Qy	958	CGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGGGAGG	1017
Db	1047	CGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGGGAGG	1106
Qy	1018	GACACCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1077
Db	1107	GACACCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1166

Qу	1078	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC	1137
Db	1167	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC	1226
Qy	1138	ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1197
Db	1227		1286
Qу	1198	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1257
Db	1287	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1346
Qy	1258	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCC	1317
Db	1347		1403
Qy	1318	CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1377
Db	1404	CGGCAGGATGGGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1463
Qy	1378	GGTGGCGCCCCCCCCCCCCCCCCCCCCCCGAGGCCGAGGAGTTCGTC	1437
Db	1464	GGTGGCGGCCACACACTGCACCACACTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1523
Qу	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1497
Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1583
Qу	1498	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1557
Db	1584	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1643
Qу	1558	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1617
Db	1644	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1703
Qу	1618	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1677
Db	1704	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1763
Qу	1678	AGCTGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1737
Db	1764	AGCTGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1823
Qу	1738	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1797
Db	1824	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1883
Qу	1798	TGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1854
Db	1884	TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1943
Qу	1855	CTGGAGGCCAGTGCCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGA	1914
Db	1944	CTGGAGGCCAGTGCCTGCTACATCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	2003

.

Qу	1915	GAGGCCCTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 1974
Db	2004	GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063
Qу	1975	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC
Db	2064	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC
Qу	2035	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2094
Db	2124	
Qy	2095	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC
Db	2184	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC
Qy	2155	CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2214
Db	2244	CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303
Qу	2215	GGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274
Db	2304	GGCACGCAGCGGTACTTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363
Qy	2275	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334
Db	2364	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423
Qy	2335	AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
Db	2424	AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
Qy	2395	GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454
Db	2484	GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543
Qy		ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
Db		ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603
Qу	2515	AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATG 2574
Db	2604	AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATG 2663
Qy	2575	ATCCTCAACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC
Db	2664	ATCCTCAACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC
Qу	2635	GCAGTGGCTGGACTGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
Db	2724	GCAGTGGCTGGACTGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
Qу	2695	TGA 2697
	2794	TGA 2786

```
RESULT 9
ADH71637
     ADH71637 standard; DNA; 2881 BP.
ID
XX
AC
     ADH71637;
XX
DT
     25-MAR-2004
                  (first entry)
XX
DE
     Human gene of the invention NOV210 SEQ ID NO:533.
XX
KW
     ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW
     anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
     vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
     obesity; diabetes; infectious disease; metabolic syndrome X;
KW
KW
     dyslipidaemia.
XX
OS
    Homo sapiens.
XX
PN
    WO2003102155-A2.
XX
PD
     11-DEC-2003.
XX
PF
     03-JUN-2003; 2003WO-US017430.
XX
     03-JUN-2002; 2002US-0385120P.
PR
PR
     04-JUN-2002; 2002US-0385784P.
PR
     05-JUN-2002; 2002US-0386041P.
PR
     05-JUN-2002; 2002US-0386047P.
PR
     06-JUN-2002; 2002US-0386376P.
     06-JUN-2002; 2002US-0386453P.
PR
     06-JUN-2002; 2002US-0386864P.
PR
     06-JUN-2002; 2002US-0387016P.
PR
     07-JUN-2002; 2002US-0386796P.
PR
     07-JUN-2002; 2002US-0386816P.
PR
PR
     07-JUN-2002; 2002US-0386931P.
PR
     07-JUN-2002; 2002US-0386942P.
     07-JUN-2002; 2002US-0386971P.
PR
     07-JUN-2002; 2002US-0387262P.
PR
PR
     08-JUN-2002; 2002US-0296960P.
PR
     10-JUN-2002; 2002US-0387400P.
PR
     10-JUN-2002; 2002US-0387535P.
     11-JUN-2002; 2002US-0387610P.
PR
     11-JUN-2002; 2002US-0387625P.
PR
     11-JUN-2002; 2002US-0387634P.
PR
     11-JUN-2002; 2002US-0387668P.
PR
     11-JUN-2002; 2002US-0387696P.
PR
     11-JUN-2002; 2002US-0387702P.
PR
PR
     11-JUN-2002; 2002US-0387836P.
PR
     11-JUN-2002; 2002US-0387859P.
     12-JUN-2002; 2002US-0387933P.
PR
     12-JUN-2002; 2002US-0387934P.
PR
     12-JUN-2002; 2002US-0387960P.
PR
     12-JUN-2002; 2002US-0388022P.
PR
     12-JUN-2002; 2002US-0388096P.
PR
PR
     13-JUN-2002; 2002US-0389123P.
PR
     14-JUN-2002; 2002US-0389118P.
     14-JUN-2002; 2002US-0389120P.
PR
```

```
PR
     14-JUN-2002; 2002US-0389144P.
     14-JUN-2002; 2002US-0389146P.
PR
PR
     17-JUN-2002; 2002US-0389729P.
PR
     17-JUN-2002; 2002US-0389742P.
     18-JUN-2002; 2002US-0389884P.
PR
PR
     19-JUN-2002; 2002US-0390006P.
     19-JUN-2002; 2002US-0390209P.
PR
     21-JUN-2002; 2002US-0390763P.
PR
     17-JUL-2002; 2002US-0396706P.
PR
PR
     06-AUG-2002; 2002US-0401628P.
     09-AUG-2002; 2002US-0402156P.
PR
PR
     09-AUG-2002; 2002US-0402256P.
PR
     09-AUG-2002; 2002US-0402389P.
     12-AUG-2002; 2002US-0402786P.
PR
     12-AUG-2002; 2002US-0402816P.
PR
     12-AUG-2002; 2002US-0402821P.
PR
     12-AUG-2002; 2002US-0402832P.
PR
     13-AUG-2002; 2002US-0403448P.
PR
     13-AUG-2002; 2002US-0403459P.
PR
     13-AUG-2002; 2002US-0403531P.
PR
     13-AUG-2002; 2002US-0403532P.
PR
     13-AUG-2002; 2002US-0403563P.
PR
     13-AUG-2002; 2002US-0406317P.
PR
    15-AUG-2002; 2002US-0403617P.
PR
     26-AUG-2002; 2002US-0406182P.
PR
     26-AUG-2002; 2002US-0406355P.
PR
PR
     27-AUG-2002; 2002US-0406240P.
     12-SEP-2002; 2002US-0410084P.
PR
     20-SEP-2002; 2002US-0412528P.
PR
     23-SEP-2002; 2002US-0412731P.
PR
     30-SEP-2002; 2002US-0414801P.
PR
     30-SEP-2002; 2002US-0414839P.
PR
     30-SEP-2002; 2002US-0414840P.
PR
PR
     30-SEP-2002; 2002US-0414954P.
PR
     09-OCT-2002; 2002US-0417186P.
     09-OCT-2002; 2002US-0417406P.
PR
     23-OCT-2002; 2002US-0420639P.
PR
     28-OCT-2002; 2002US-0421156P.
PR
     31-OCT-2002; 2002US-0422690P.
PR
     01-NOV-2002; 2002US-0423130P.
PR
     05-NOV-2002; 2002US-00423798.
PR
PR
     05-NOV-2002; 2002US-0423798P.
PR
     12-NOV-2002; 2002US-0425453P.
XX
PA
     (CURA-) CURAGEN CORP.
XX
PΙ
    Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PΙ
     Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PΙ
     Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E,
                                                                     Guo X;
PΙ
     Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PΙ
    Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
     Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PΙ
     Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PΙ
PΙ
     Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI
     Zhong H;
XX
DR
     WPI; 2004-081935/08.
```

```
DR
    P-PSDB; ADH71638.
XX
PT
    New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT
    treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
    obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
PT
XX
    Example 21; SEQ ID NO 533; 1880pp; English.
PS
XX
CC
    The invention relates to a novel isolated polypeptide (NOVX). A
CC
    polypeptide of the invention has cytostatic, immunomodulator,
    neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC
    antilipaemic activity, and may have a use in gene therapy, and as a
CC
CC
    vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC
    any of the 303 fully defined nucleotide sequences given in the
    specification. The polypeptide is useful in the manufacture of a
CC
CC
    medicament for treating a syndrome associated with a human disease. The
    polypeptide, polynucleotide and antibody are useful in diagnosing,
CC
    treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC
    Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC
    diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC
CC
    further used as hybridisation probes, in chromosome mapping, tissue
    typing, preventive medicine, and pharmacogenomics. The present sequence
CC
CC
    encodes a NOVX polypeptide of the invention.
XX
    Sequence 2881 BP; 527 A; 985 C; 867 G; 502 T; 0 U; 0 Other;
SO
 Query Match
                      97.1%;
                             Score 2619.8;
                                          DB 12; Length 2881;
 Best Local Similarity
                      98.9%;
                             Pred. No. 0;
 Matches 2672; Conservative
                            0; Mismatches
                                               Indels
                                                        9;
                                                           Gaps
                                                                   3;
                                           22;
          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
Qy
            87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146
Db
         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
Qу
            147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206 :
Db
        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
Qу
            207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266
Db
        181 GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
Qy
            267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326
Db
        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCAGACGGGAGCAGTGGGCTGCCC 300
Qу
            327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386
Db
        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
Qу
            387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446
Db
        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
Qу
```

447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506

Db

Qу	421	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
Db .	507	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	566
Qу	481	TCCCTGGAGCAGGGCATCGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Db	567	TCCCTGGAGCAGGGCATCGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	626
Qy	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	600
Db	627	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATC	686
Qу	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
Db	68.7	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	746
Qу	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTAC	720
Db	747	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTAC	806
Qу	721	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	780
Db	807	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCC	866
Qу	781	GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
Db	867	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	926
Qу	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC	897
Db	927	TGTGAGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGTGTCTCTGTGGACGGC	986
Qу	898	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	957
Db	987	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	1046
Qy	958	CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG	1017
Db	1047	CGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGGGAGG	1106
Qy	1018	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1077
Db	1107	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1166
Qy	1078	GCCCTCTATGTGGGCCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC	1137
Db	1167	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTTGTCCTC	1226
Qу	1138	ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1197
Db	1227	ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCCATT	1286
Qy	1198	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCAGCAAAGCAGACAACCCCCATCTG	1257
Db	1287	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1346

	QУ	1258	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCC	1317
	Db	1347	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCC	1403
	Qy	1318	CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1377
	Db	1404	CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1463
	Qу	1378	GGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1437
	Db	1464	GGTGGCGGCCACACACTGCACCACACTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1523
	Qу	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1497
	Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1583
•	Qy	1498	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1557
	Db	1584		1643
	Qу	1558	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1617
	Db	1644		1703
	Qу	1618	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1677
	Db	1704		1763
	Qу	1678	AGCTGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1737
	Db	1764		1823
	Qу	1738	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1797
	Db	1824	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1883
	Qу	1798	TGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1854
	Db	1884	TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1943
	Qу	1855	$\tt CTGGAGGCCAGTGCCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA$	1914
	Db	1944	CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	2003
	Qу	1915	GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	1974
	Db	2004	GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	2063
	Qу	1975	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2034
	Db	2064	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2123
	Qу	2035	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC	2094
	Db	2124	AAGGAGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC	2183
	Ωv	2095	C T	215/

ί,

```
Db
      2155 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2214
Qу
         2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303
Db
      2215 GGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274
Qy
         2304 GGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363
Db
      2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGCGACGGGCAGAGCTTCAGCATC 2334
Qу
         2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423
Db
      2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
Qy
         2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
Db
      2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454
Qу
         2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543
Db
      2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
Qy
         2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603
Db
      2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2574
Qу
         2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663
Db
      Qу
         Db
      2635 GCAGTGGCTGGACTGGCCAGCCAGACGCTGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
Qy
         2724 GCAGTGGCTGGACTGGCCAGCCAGACGCTGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
Db
      2695 TGA 2697
Qу
        . 111
Db
      2784 TGA 2786
RESULT 10
ADH71641
ID
   ADH71641 standard; DNA; 2881 BP.
XX
AC
   ADH71641;
XX
DT
   25-MAR-2004 (first entry)
XX
DE
   Human gene of the invention NOV21g SEQ ID NO:537.
XX
KW
   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW
   anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
   vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
```

KW

```
KW
     obesity; diabetes; infectious disease; metabolic syndrome X;
KW
     dyslipidaemia.
XX
os
     Homo sapiens.
XX
     WO2003102155-A2.
PN
XX
PD
     11-DEC-2003.
XX
PF
     03-JUN-2003; 2003WO-US017430.
XX
PR
     03-JUN-2002; 2002US-0385120P.
     04-JUN-2002; 2002US-0385784P.
PR
     05-JUN-2002; 2002US-0386041P.
PR
     05-JUN-2002; 2002US-0386047P.
PR
     06-JUN-2002; 2002US-0386376P.
PR
     06-JUN-2002; 2002US-0386453P.
PR
     06-JUN-2002; 2002US-0386864P.
PR
     06-JUN-2002; 2002US-0387016P.
PR
     07-JUN-2002; 2002US-0386796P.
PR
     07-JUN-2002; 2002US-0386816P.
PR
     07-JUN-2002; 2002US-0386931P.
PR
     07-JUN-2002; 2002US-0386942P.
PR
     07-JUN-2002; 2002US-0386971P.
PR
     07-JUN-2002; 2002US-0387262P.
PR
PR
     08-JUN-2002; 2002US-0296960P.
PR
     10-JUN-2002; 2002US-0387400P.
PR
     10-JUN-2002; 2002US-0387535P.
PR
     11-JUN-2002; 2002US-0387610P.
     11-JUN-2002; 2002US-0387625P.
PR
     11-JUN-2002; 2002US-0387634P.
PR
     11-JUN-2002; 2002US-0387668P.
PR
     11-JUN-2002; 2002US-0387696P.
PR
     11-JUN-2002; 2002US-0387702P.
PR
     11-JUN-2002; 2002US-0387836P.
PR
     11-JUN-2002; 2002US-0387859P.
PR
     12-JUN-2002; 2002US-0387933P.
PR
     12-JUN-2002; 2002US-0387934P.
PR
PR
     12-JUN-2002; 2002US-0387960P.
PR
     12-JUN-2002; 2002US-0388022P.
PR
     12-JUN-2002; 2002US-0388096P.
PR
     13-JUN-2002; 2002US-0389123P.
     14-JUN-2002; 2002US-0389118P.
PR
     14-JUN-2002; 2002US-0389120P.
PR
     14-JUN-2002; 2002US-0389144P.
PR
     14-JUN-2002; 2002US-0389146P.
PR
     17-JUN-2002; 2002US-0389729P.
PR
PR
     17-JUN-2002; 2002US-0389742P.
PR
     18-JUN-2002; 2002US-0389884P.
     19-JUN-2002; 2002US-0390006P.
PR
     19-JUN-2002; 2002US-0390209P.
PR
     21-JUN-2002; 2002US-0390763P.
PR
PR
     17-JUL-2002; 2002US-0396706P.
PR
     06-AUG-2002; 2002US-0401628P.
PR
     09-AUG-2002; 2002US-0402156P.
PR
     09-AUG-2002; 2002US-0402256P.
PR
     09-AUG-2002; 2002US-0402389P.
```

```
12-AUG-2002; 2002US-0402786P.
PR
     12-AUG-2002; 2002US-0402816P.
PR
     12-AUG-2002; 2002US-0402821P.
PR
     12-AUG-2002; 2002US-0402832P.
PR
     13-AUG-2002; 2002US-0403448P.
PR
     13-AUG-2002; 2002US-0403459P.
PR
     13-AUG-2002; 2002US-0403531P.
PR
     13-AUG-2002; 2002US-0403532P.
PR
     13-AUG-2002; 2002US-0403563P.
PR
     13-AUG-2002; 2002US-0406317P.
PR
     15-AUG-2002; 2002US-0403617P.
PR
PR
     26-AUG-2002; 2002US-0406182P.
     26-AUG-2002; 2002US-0406355P.
PR
     27-AUG-2002; 2002US-0406240P.
PR
     12-SEP-2002; 2002US-0410084P.
PR
     20-SEP-2002; 2002US-0412528P.
PR
PR
     23-SEP-2002; 2002US-0412731P.
     30-SEP-2002; 2002US-0414801P.
PR
PR
     30-SEP-2002; 2002US-0414839P.
     30-SEP-2002; 2002US-0414840P.
PR
     30-SEP-2002; 2002US-0414954P.
PR
     09-OCT-2002; 2002US-0417186P.
PR
PR
     09-OCT-2002; 2002US-0417406P.
     23-OCT-2002; 2002US-0420639P.
PR
     28-OCT-2002; 2002US-0421156P.
PR
PR
     31-OCT-2002; 2002US-0422690P.
     01-NOV-2002; 2002US-0423130P.
PR
PR
     05-NOV-2002; 2002US-00423798.
     05-NOV-2002; 2002US-0423798P.
PR
     12-NOV-2002; 2002US-0425453P.
PR
XX
PA
     (CURA-) CURAGEN CORP.
XX
PΙ
    Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
     Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PΙ
PΙ
     Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
     Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PΙ
    Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PΙ
     Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PΙ
PΙ
                                 Sciore P, Shenoy SG,
     Rieger DK, Rothenberg ME,
                                                        Shimkets RA;
PΙ
     Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI
     Zhong H;
XX
DR
     WPI; 2004-081935/08.
DR
     P-PSDB; ADH71642.
XX
PT
    New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT
     treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT
     obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS
     Example 21; SEQ ID NO 537; 1880pp; English.
XX
CC
     The invention relates to a novel isolated polypeptide (NOVX). A
CC
     polypeptide of the invention has cytostatic, immunomodulator,
CC
     neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC
     antilipaemic activity, and may have a use in gene therapy, and as a
CC
     vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
```

CC any of the 303 fully defined nucleotide sequences given in the CC specification. The polypeptide is useful in the manufacture of a CC medicament for treating a syndrome associated with a húman disease. The CC polypeptide, polynucleotide and antibody are useful in diagnosing, treating or preventing NOVX-associated disorders, e.g. cancer, cachexia, CC CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are CC further used as hybridisation probes, in chromosome mapping, tissue CC typing, preventive medicine, and pharmacogenomics. The present sequence CC CC encodes a NOVX polypeptide of the invention. XX

SQ

Sequence 2881 BP; 526 A; 986 C; 867 G; 502 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2881; Best Local Similarity 98.9%; Pred. No. 0; Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

		•	
Qу	1	ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC	60
Db	87	ATGGCCGTCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC	146
Qy	61	CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGCCCAACCCG	120
Db	147	CGCGGCTCGGGTGCCCAGCAGAGTGCCACCCGTGCCCAACCCG	206
Qу	121	GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA	180
Db	207	GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA	266
Qy	181	GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG	240
Db	267	GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG	326
Qу		TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC	
Db		TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG	
Qy .		ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG	
Db		ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG	
Qу		GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	
Db		GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	
Qy		TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	
Db .		TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	
Qy		TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	
Db		TCCCTGGAGCAGGGCATCGTGCCGTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	
Qy		GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATC	
Db	627	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATC	686

	QУ	601	ACGUGGGAGUAUAGCUTGGTGGTGUGACAGGCUCGCUTTGUTGAUAUGGUUAAUTAUACU	660
	Db	687		746
	Qy .	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTCATCGTCTAC	720
	Db	747	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTCATCGTCTAC	806
	Qу	721	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCC	780
	Db	807		866
	Qу	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
	Db	867		926
	Qу	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC	897
	Db	927	TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTGTGTCTCTGTGGACGGC	986
	Qу	898	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	957
	Db	987	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	1046
	Qу	958	CGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGGGAGG	1017
	Db	1047	CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG	1106
	Qу	1018	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1077
	Db	1107	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1166
	QУ	1078	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC	1137
	Db	1167	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC	1226
·	Qу		ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1197
	Db		ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1286
	Qу	1198	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1257
	Db	1287	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1346
	QУ	1258	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCC	1317
	Db	1347	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCC	1403
	Qу	1318	CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1377
	Db	1404	CGGCAGGATGGGCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1463
	Qy	1378	GGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1437
	Db	1464		1523
	Ov	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1497

,

Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1583
Qу	1498	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1557
Db	1584	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1643
Qу	1558	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1617
Db	1644	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1703
Qу	1618	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1677
Db	1704	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1763
Qу	1678	AGCTGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1737
Db	1764	AGCTGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1823
Qу	1738	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1797
Db	1824	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1883
Qу	1798	TGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1854
Db	1884	TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1943
Qу	1855	CTGGAGGCCAGTGCCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGA	1914
Db	1944	CTGGAGGCCAGTGCCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	2003
Qу	1915	GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	1974
Db	2004	GAGGCCCTCAGCGTGGCTGCCCCAAGCCCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	2063
Qy	1975	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2034
Db	2064	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2123
.Qy	2035	AAGGAGGTGCTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC	2094
Db	2124	AAGGAGCTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC	2183
Qу	2095	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	2154
Db	2184	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCACGATGTGCCCAGCTCC	2243
Qy	2155	CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT	2214
Db	2244	CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT	2303
Qy	2215	GGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT	2274
Db	2304	GGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT	2363
Qy	2275	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC	2334

```
Db
       2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGCGACGGGCAGAGCTTCAGCATC 2423
       2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
Qy
          2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
Db
       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454
Qу
          2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543
Db
       2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
Qу
          2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603
Db
       2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATG 2574
Qy
          2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATG 2663
Db
       Qy
          Db
       2635 GCAGTGGCTGGACTGGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
Qy
          2724 GCAGTGGCTGGACTGGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
Dh
       2695 TGA 2697
Qу
          111
       2784 TGA 2786
Dh
RESULT 11
ADH71629
ID
   ADH71629 standard; DNA; 2881 BP.
XX
AC
   ADH71629;
XX
DT
   25-MAR-2004 (first entry)
XX
DE
   Human gene of the invention NOV21k SEQ ID NO:525.
XX
   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW
KW
   anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW
   vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
   obesity; diabetes; infectious disease; metabolic syndrome X;
KW
   dyslipidaemia.
XX
os
   Homo sapiens.
XX
PN
   WO2003102155-A2.
XX
PD
   11-DEC-2003.
XX
   03-JUN-2003; 2003WO-US017430.
PF
XX
PR
   03-JUN-2002; 2002US-0385120P.
    04-JUN-2002; 2002US-0385784P.
PR
```

```
PR
     05-JUN-2002; 2002US-0386041P.
PR
     05-JUN-2002; 2002US-0386047P.
PR
     06-JUN-2002; 2002US-0386376P.
     06-JUN-2002; 2002US-0386453P.
PR
PR
     06-JUN-2002; 2002US-0386864P.
PR
     06-JUN-2002; 2002US-0387016P.
     07-JUN-2002; 2002US-0386796P.
PR
     07-JUN-2002; 2002US-0386816P.
PR
PR
     07-JUN-2002; 2002US-0386931P.
PR
     07-JUN-2002; 2002US-0386942P.
     07-JUN-2002; 2002US-0386971P.
PR
     07-JUN-2002; 2002US-0387262P.
PR
     08-JUN-2002; 2002US-0296960P.
PR
     10-JUN-2002; 2002US-0387400P.
PR
     10-JUN-2002; 2002US-0387535P.
PR
     11-JUN-2002; 2002US-0387610P.
PR
     11-JUN-2002; 2002US-0387625P.
PR
     11-JUN-2002; 2002US-0387634P.
PR
     11-JUN-2002; 2002US-0387668P.
PR
     11-JUN-2002; 2002US-0387696P.
PR
     11-JUN-2002; 2002US-0387702P.
PR
     11-JUN-2002; 2002US-0387836P.
PR
     11-JUN-2002; 2002US-0387859P.
PR
     12-JUN-2002; 2002US-0387933P.
PR
     12-JUN-2002; 2002US-0387934P.
PR
     12-JUN-2002; 2002US-0387960P.
PR
PR
     12-JUN-2002; 2002US-0388022P.
PR
     12-JUN-2002; 2002US-0388096P.
PR
     13-JUN-2002; 2002US-0389123P.
     14-JUN-2002; 2002US-0389118P.
PR
     14-JUN-2002; 2002US-0389120P.
PR
     14-JUN-2002; 2002US-0389144P.
PR
     14-JUN-2002; 2002US-0389146P.
PR
PR
     17-JUN-2002; 2002US-0389729P.
PR
     17-JUN-2002; 2002US-0389742P.
     18-JUN-2002; 2002US-0389884P.
PR
     19-JUN-2002; 2002US-0390006P.
PR
     19-JUN-2002; 2002US-0390209P.
PR
PR
     21-JUN-2002; 2002US-0390763P.
     17-JUL-2002; 2002US-0396706P.
PR
     06-AUG-2002; 2002US-0401628P.
PR
     09-AUG-2002; 2002US-0402156P.
PR
     09-AUG-2002; 2002US-0402256P.
PR
     09-AUG-2002; 2002US-0402389P.
PR
     12-AUG-2002; 2002US-0402786P.
PR
     12-AUG-2002; 2002US-0402816P.
PR
     12-AUG-2002; 2002US-0402821P.
PR
     12-AUG-2002; 2002US-0402832P.
PR
PR
     13-AUG-2002; 2002US-0403448P.
PR
     13-AUG-2002; 2002US-0403459P.
PR
     13-AUG-2002; 2002US-0403531P.
     13-AUG-2002; 2002US-0403532P.
PR
     13-AUG-2002; 2002US-0403563P.
PR
     13-AUG-2002; 2002US-0406317P.
PR
     15-AUG-2002; 2002US-0403617P.
PR
PR
     26-AUG-2002; 2002US-0406182P.
     26-AUG-2002; 2002US-0406355P.
PR
```

```
27-AUG-2002; 2002US-0406240P.
PR
     12-SEP-2002; 2002US-0410084P.
PR
     20-SEP-2002; 2002US-0412528P.
PR
     23-SEP-2002; 2002US-0412731P.
PR
     30-SEP-2002; 2002US-0414801P.
PR
     30-SEP-2002; 2002US-0414839P.
PR
     30-SEP-2002; 2002US-0414840P.
PR
     30-SEP-2002; 2002US-0414954P.
PR
     09-OCT-2002; 2002US-0417186P.
PR
     09-OCT-2002; 2002US-0417406P.
PR
     23-OCT-2002; 2002US-0420639P.
PR
· PR
     28-OCT-2002; 2002US-0421156P.
     31-OCT-2002; 2002US-0422690P.
PR
     01-NOV-2002; 2002US-0423130P.
PR
     05-NOV-2002; 2002US-00423798.
PR
PR
     05-NOV-2002; 2002US-0423798P.
PR
     12-NOV-2002; 2002US-0425453P.
XX
PA
     (CURA-) CURAGEN CORP.
XX
     Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI
     Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI
     Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PΙ
     Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PΙ
     Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI
                                                                 Rastelli L;
PI
     Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D,
PΙ
     Rieger DK, Rothenberg ME,
                                 Sciore P, Shenoy SG,
                                                        Shimkets RA;
PΙ
                  Spytek KA, Stone DJ, Vernet CAM, Voss EZ,
     Smithson G,
PΙ
     Zhong H;
XX
DR
     WPI; 2004-081935/08.
DR
     P-PSDB; ADH71630.
XX
     New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT
PΤ
     treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT
     obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS
     Example 21; SEQ ID NO 525; 1880pp; English.
XX
CC
     The invention relates to a novel isolated polypeptide (NOVX). A
     polypeptide of the invention has cytostatic, immunomodulator,
CC
CC
     neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC
     antilipaemic activity, and may have a use in gene therapy, and as a
CC
     vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC
     any of the 303 fully defined nucleotide sequences given in the
CC
     specification. The polypeptide is useful in the manufacture of a
     medicament for treating a syndrome associated with a human disease. The
CC
     polypeptide, polynucleotide and antibody are useful in diagnosing,
CC
     treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC
CC
     Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC
     diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
     further used as hybridisation probes, in chromosome mapping, tissue
CC
CC
     typing, preventive medicine, and pharmacogenomics. The present sequence
CC
     encodes a NOVX polypeptide of the invention.
XX
     Sequence 2881 BP; 526 A; 986 C; 868 G; 501 T; 0 U; 0 Other;
SQ
```

Query Match 97.1%; Score 2619.8; DB 12; Length 2881; Best Local Similarity 98.9%; Pred. No. 0; Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

Macches	2012	z, conservative o, mismatches zz, inders 9, daps	٠,
Qy	1	ATGGCCGTCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60	ı
Db	87	ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 14	6
Qу	61	CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 12	0
Db	147	CGCGGCTCGGGTGCCCAGCAGAGTGCCACCCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 20	6
Qy	121	GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 18	0
Db	207	GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 26	6
Qy	181	GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 24	0
Db	267	GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 32	6
Qу	241	TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 30	0
Db	327	TGGGTGCGCCAGGTGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 38	6
QУ	301	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 36	0
Db	387	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 44	6
Qу	361	GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 42	0
Db	447	GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 50	6
Qу		TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 48	
Db		TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 56	
Qу		TCCCTGGAGCAGGGCATCGTGCCGTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	
Db		TCCCTGGAGCAGGGCATCGTGCCGTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 62	
Qу		GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 60	
Db	627	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATC 68	6
Qу		ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 66	
Db	687	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 74	6
Qу	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 72	0
Db	747	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 80	6
Qу		GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 78	
Db	807	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 86	6
QУ	781	GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 84	0

Db	867	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	926
Qу	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC	897
Db	927	TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC	986
Qу	898	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	957
Db	987	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	1046
Qу	958	CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTG	1017
Db	1047		1106
Qу	1018	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1077
Db	1107		1166
Qу	1078	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC	1137
Db	1167		1226
Qу	1138	ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1197
Db	1227		1286
Qу	1198	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCAGCAAAGCAGACAACCCCCATCTG	1257
Db	1287	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1346
Qу	1258	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCAGGGCAGTCTCTGTCCC	1317
Db	1347	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCC	1403
Qу	1318	CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1377
Db	1404	CGGCAGGATGGGCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1463
Qу	1378	GGTGGCGCCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1437
Db	1464	GGTGGCGGCCACACACTGCACCACACCTCTCAGGCCGAGGAGTTCGTC	1523
Qу	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1497
Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1583
Qу	1498	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1557
Db	1584	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1643
Qy	1558	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1617
Db	1644		1703
Qу	1618	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1677

.

Db	1704	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1763
Qy	1678	AGCTGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1737
Db	1764	AGCTGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1823
Qу	1738	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1797
Db	1824	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1883
Qу	1798	TGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1854
Db	1884	TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1943
Qу	1855	CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	1914
Db	1944	CTGGAGGCCAGTGCCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	2003
Qу	1915	GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	1974
Db	2004	GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	2063
Qу	1975	TGCACCTCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC	2034
Db	2064	TGCACCTCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC	2123
QУ	2035	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC	2094
Db	2124	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC	2183
Qу	2095	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	2154
Db	2184	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	2243
Qу	2155	CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT	2214
Db	2244	CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT	2303
Qу	2215	GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT	2274
Db	2304	GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT	2363
Qу	2275	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC	2334
Db	2364	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC	2423
QУ	2335	AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2394
Db		AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	
Qу		GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG	
Db		GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG	
Qу		ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG	
Db	2544	ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG	2603

```
Qу
       2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATG 2574
           2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663
Db
       Qy
           Db
       2635 GCAGTGGCTGGACTGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
Qу
           2724 GCAGTGGCTGGACTGGCCAGCCAGACGCTGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
Db
       2695 TGA 2697
Qу
           2784 TGA 2786
Db
RESULT 12
ADH71631
    ADH71631 standard; DNA; 2881 BP.
ID
XX
AC
    ADH71631;
XX
    25-MAR-2004 (first entry)
DT
XX
DE
    Human gene of the invention NOV211 SEQ ID NO:527.
XX
KW
    ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW
    anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW
    vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
    obesity; diabetes; infectious disease; metabolic syndrome X;
KW
KW
    dyslipidaemia.
XX
os
    Homo sapiens.
XX
    WO2003102155-A2.
PN
XX
PD
    11-DEC-2003.
XX
    03-JUN-2003; 2003WO-US017430.
PF
XX
PR
    03-JUN-2002; 2002US-0385120P.
PR
    04-JUN-2002; 2002US-0385784P.
    05-JUN-2002; 2002US-0386041P.
PR
    05-JUN-2002; 2002US-0386047P.
PR
    06-JUN-2002; 2002US-0386376P.
PR
    06-JUN-2002; 2002US-0386453P.
PR
    06-JUN-2002; 2002US-0386864P.
PR
    06-JUN-2002; 2002US-0387016P.
PR
    07-JUN-2002; 2002US-0386796P.
PR
    07-JUN-2002; 2002US-0386816P.
PR
    07-JUN-2002; 2002US-0386931P.
PR
    07-JUN-2002; 2002US-0386942P.
PR
PR
    07-JUN-2002; 2002US-0386971P.
PR
    07-JUN-2002; 2002US-0387262P.
    08-JUN-2002; 2002US-0296960P.
PR
```

```
10-JUN-2002; 2002US-0387400P.
PR
     10-JUN-2002; 2002US-0387535P.
PR
PR
     11-JUN-2002; 2002US-0387610P.
     11-JUN-2002; 2002US-0387625P.
PR
     11-JUN-2002; 2002US-0387634P.
PR
     11-JUN-2002; 2002US-0387668P.
PR
     11-JUN-2002; 2002US-0387696P.
PR
     11-JUN-2002; 2002US-0387702P.
PR
PR
     11-JUN-2002; 2002US-0387836P.
PR
     11-JUN-2002; 2002US-0387859P.
     12-JUN-2002; 2002US-0387933P.
PR
PR
     12-JUN-2002; 2002US-0387934P.
     12-JUN-2002; 2002US-0387960P.
PR
     12-JUN-2002; 2002US-0388022P.
PR
     12-JUN-2002; 2002US-0388096P.
PR
     13-JUN-2002; 2002US-0389123P.
PR
PR
     14-JUN-2002; 2002US-0389118P.
     14-JUN-2002; 2002US-0389120P.
PR
PR
     14-JUN-2002; 2002US-0389144P.
     14-JUN-2002; 2002US-0389146P.
PR
     17-JUN-2002; 2002US-0389729P.
PR
     17-JUN-2002; 2002US-0389742P.
PR
     18-JUN-2002; 2002US-0389884P.
PR
PR
     19-JUN-2002; 2002US-0390006P.
     19-JUN-2002; 2002US-0390209P.
PR
PR
     21-JUN-2002; 2002US-0390763P.
PR
     17-JUL-2002; 2002US-0396706P.
     06-AUG-2002; 2002US-0401628P.
PR
     09-AUG-2002; 2002US-0402156P.
PR
     09-AUG-2002; 2002US-0402256P.
PR
PR
     09-AUG-2002; 2002US-0402389P.
PR
     12-AUG-2002; 2002US-0402786P.
PR
     12-AUG-2002; 2002US-0402816P.
PR
     12-AUG-2002; 2002US-0402821P.
     12-AUG-2002; 2002US-0402832P.
PR
     13-AUG-2002; 2002US-0403448P.
PR
     13-AUG-2002; 2002US-0403459P.
PR
     13-AUG-2002; 2002US-0403531P.
PR
     13-AUG-2002; 2002US-0403532P.
PR
PR
     13-AUG-2002; 2002US-0403563P.
PR
     13-AUG-2002; 2002US-0406317P.
PR
     15-AUG-2002; 2002US-0403617P.
     26-AUG-2002; 2002US-0406182P.
PR
     26-AUG-2002; 2002US-0406355P.
PR
     27-AUG-2002; 2002US-0406240P.
PR
PR
     12-SEP-2002; 2002US-0410084P.
PR
     20-SEP-2002; 2002US-0412528P.
PR
     23-SEP-2002; 2002US-0412731P.
PR
     30-SEP-2002; 2002US-0414801P.
     30-SEP-2002; 2002US-0414839P.
PR
     30-SEP-2002; 2002US-0414840P.
PR
     30-SEP-2002; 2002US-0414954P.
PR
PR
     09-OCT-2002; 2002US-0417186P.
     09-OCT-2002; 2002US-0417406P.
PR
     23-OCT-2002; 2002US-0420639P.
PR
PR
     28-OCT-2002; 2002US-0421156P.
     31-OCT-2002; 2002US-0422690P.
PR
```

```
PR
     01-NOV-2002; 2002US-0423130P.
     05-NOV-2002; 2002US-00423798.
PR
     05-NOV-2002; 2002US-0423798P.
PR
PR
    12-NOV-2002; 2002US-0425453P.
XX
PA
     (CURA-) CURAGEN CORP.
XX
PΙ
    Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PΙ
    Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PΙ
    Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E,
PΙ
    Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PΙ
    Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
    Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PΙ
ΡI
    Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
    Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PΙ
PΙ
    Zhong H;
XX
DR
    WPI; 2004-081935/08.
DR
    P-PSDB; ADH71632.
XX
PT
    New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT
    treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT
    obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS
    Example 21; SEQ ID NO 527; 1880pp; English.
XX
CC
    The invention relates to a novel isolated polypeptide (NOVX). A
CC
    polypeptide of the invention has cytostatic, immunomodulator,
    neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC
CC
    antilipaemic activity, and may have a use in gene therapy, and as a
CC
    vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
    any of the 303 fully defined nucleotide sequences given in the
CC
CC
    specification. The polypeptide is useful in the manufacture of a
CC
    medicament for treating a syndrome associated with a human disease. The
CC
    polypeptide, polynucleotide and antibody are useful in diagnosing,
CC
    treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC
    Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC
    diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC
    further used as hybridisation probes, in chromosome mapping, tissue
CC.
    typing, preventive medicine, and pharmacogenomics. The present sequence
CC
    encodes a NOVX polypeptide of the invention.
XX
SO
    Sequence 2881 BP; 526 A; 984 C; 868 G; 503 T; 0 U; 0 Other;
                                Score 2619.8; DB 12; Length 2881;
 Query Match
                         97.1%;
 Best Local Similarity
                        98.98;
                                Pred. No. 0;
                                               22; Indels
 Matches 2672; Conservative
                               0; Mismatches
                                                              9; Gaps
                                                                         3;
Qу
           1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
             Db
          87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146
          61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
Qу
             Db
         147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206
Qy
         121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
```

Db	207		266
		GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG	
Qу			
Db		GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG	
Qу	241	TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC	300
Db ·	327	$\tt TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCGGGAGCAGGGAGCAGGAGCAGGAGG$	386
Qу	301	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG	360
Db	387	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG	446
Qу	361	GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	420
Db	447		506
Qу	421	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
Db	507	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	566
Qy	481	TCCCTGGAGCAGGCATCGTGCCCTGCCGTCCACCGGAGGCCATCCCTCCAGCCGAG	540
Db	567	TCCCTGGAGCAGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	626
Qу	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	600
Db	627	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	686
Qу	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
Db	687		746
Qу	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC	720
Db	747	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTAC	806
Qу	721		780
Db	807		866
Qу	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
Db	867		926
Qу	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC	897
Db	927		986
Qу	898	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	957
Db	987	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	1046
Qy	958	CGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGGGAGG	1017

Db .	1047	CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG	1106
Qy	1018	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1077
Db	1107	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1166
Qy	1078	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC	1137
Db	1167	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC	1226
Qу	1138	ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1197
Db	1227	ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1286
Qу	1198	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1257
Db	1287	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1346
Qу	1258	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCAGGGCAGTCTCTGTCCC	1317
Db	1347	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCC	1403
Qу	1318	CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1377
Db	1404	CGGCAGGATGGGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1463
Qу	1378	GGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1437
Db	1464	GGTGGCGGCCACACACTGCACCACACTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1523
Qy `	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1497
Db ·	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1583
Qу	1498	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1557
Db	1584	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1643
Qy	1558	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1617
Db	1644	CTCATCCCCCAGATGTCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1703
Qy	1618	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1677
Db	1704	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1763
Qy	1678	AGCTGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1737
Db	1764	AGCTGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1823
Qy	1738	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1797
Db	1824	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1883
Qу	1798	TGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1854
Db	1884	TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1943
		·	

.

	Qy	1855	CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	1914
	Db	1944	CTGGAGGCCAGTGCCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	2003
	Qy .	1915	GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	1974
	Db	2004	GAGGCCCTCAGCGTGGCTGCCCCAAGCCCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	2063
	Qy	1975	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2034
	Db	2064	TGCACCTCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC	2123
	Qy	2035	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC	2094
	Db	2124	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC	2183
	Qy	2095	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	2154
	Db	2184	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	2243
	Qy	2155	CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT	2214
	Db	2244	CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT	2303
	QУ	2215	GGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT	2274
	Db	2304	GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT	2363
	Qу	2275	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC	2334
	Db		GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC	
•	Qу		AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	
	Db		AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	
	Qy		GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG	
	Db		GGGGTCCCAGCCCTGGTGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG	
	Qу		ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG	
	Db		ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG	
	ДУ		AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATG	
	Db		AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATG	
	Qy		ATCCTCAACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC	
	Db		ATCCTCAACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC	
	Qy		GCAGTGGCTGGACTGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC	
	Db	2124	GCAGTGGCTGGACTGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC	2/83

```
2695 TGA 2697
Qy
              111
         2784 TGA 2786
Db
RESULT 13
ADH71645
     ADH71645 standard; DNA; 2881 BP.
ID
XX
AC
     ADH71645;
XX
DT
     25-MAR-2004
                  (first entry)
XX
DE
     Human gene of the invention NOV21s SEQ ID NO:541.
XX
KW
     ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
     anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW
KW
     vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
     obesity; diabetes; infectious disease; metabolic syndrome X;
KW
     dyslipidaemia.
XX
os
     Homo sapiens.
XX
PN
     WO2003102155-A2.
XX
PD
     11-DEC-2003.
XX
     03-JUN-2003; 2003WO-US017430.
PF
XX
     03-JUN-2002; 2002US-0385120P.
PR
PR
     04-JUN-2002; 2002US-0385784P.
PR
     05-JUN-2002; 2002US-0386041P.
PR
     05-JUN-2002; 2002US-0386047P.
PR
     06-JUN-2002; 2002US-0386376P.
PR
     06-JUN-2002; 2002US-0386453P.
PR
     06-JUN-2002; 2002US-0386864P.
     06-JUN-2002; 2002US-0387016P.
PR
     07-JUN-2002; 2002US-0386796P.
PR
     07-JUN-2002; 2002US-0386816P.
PR
PR
     07-JUN-2002; 2002US-0386931P.
PR
     07-JUN-2002; 2002US-0386942P.
     07-JUN-2002; 2002US-0386971P.
PR
     07-JUN-2002; 2002US-0387262P.
PR
     08-JUN-2002; 2002US-0296960P.
PR
     10-JUN-2002; 2002US-0387400P.
PR
     10-JUN-2002; 2002US-0387535P.
PR
     11-JUN-2002; 2002US-0387610P.
PR
PR
     11-JUN-2002; 2002US-0387625P.
PR
     11-JUN-2002; 2002US-0387634P.
     11-JUN-2002; 2002US-0387668P.
PR
     11-JUN-2002; 2002US-0387696P.
PR
     11-JUN-2002; 2002US-0387702P.
PR
     11-JUN-2002; 2002US-0387836P.
PR
     11-JUN-2002; 2002US-0387859P.
PR
PR
     12-JUN-2002; 2002US-0387933P.
PR
     12-JUN-2002; 2002US-0387934P.
```

PR

12-JUN-2002; 2002US-0387960P.

```
PR
     12-JUN-2002; 2002US-0388022P.
     12-JUN-2002; 2002US-0388096P.
PR
PR
     13-JUN-2002; 2002US-0389123P.
PR
     14-JUN-2002; 2002US-0389118P.
     14-JUN-2002; 2002US-0389120P.
PR
PR
     14-JUN-2002; 2002US-0389144P.
     14-JUN-2002; 2002US-0389146P.
PR
     17-JUN-2002; 2002US-0389729P.
PR
PR
     17-JUN-2002; 2002US-0389742P.
     18-JUN-2002; 2002US-0389884P.
PR
     19-JUN-2002; 2002US-0390006P.
PR
PR
     19-JUN-2002; 2002US-0390209P.
PR
     21-JUN-2002; 2002US-0390763P.
     17-JUL-2002; 2002US-0396706P.
PR
     06-AUG-2002; 2002US-0401628P.
PR
     09-AUG-2002; 2002US-0402156P.
PR
     09-AUG-2002; 2002US-0402256P.
PR
     09-AUG-2002; 2002US-0402389P.
PR
     12-AUG-2002; 2002US-0402786P.
PR
PR
     12-AUG-2002; 2002US-0402816P.
PR
     12-AUG-2002; 2002US-0402821P.
     12-AUG-2002; 2002US-0402832P.
PR
     13-AUG-2002; 2002US-0403448P.
PR
     13-AUG-2002; 2002US-0403459P.
PR
PR
     13-AUG-2002; 2002US-0403531P.
     13-AUG-2002; 2002US-0403532P.
PR
PR
     13-AUG-2002; 2002US-0403563P.
PR
     13-AUG-2002; 2002US-0406317P.
PR
     15-AUG-2002; 2002US-0403617P.
     26-AUG-2002; 2002US-0406182P.
PR
     26-AUG-2002; 2002US-0406355P.
PR
PR
     27-AUG-2002; 2002US-0406240P.
PR
     12-SEP-2002; 2002US-0410084P.
PR
     20-SEP-2002; 2002US-0412528P.
PR
     23-SEP-2002; 2002US-0412731P.
     30-SEP-2002; 2002US-0414801P.
PR
     30-SEP-2002; 2002US-0414839P.
PR
     30-SEP-2002; 2002US-0414840P.
PR
     30-SEP-2002; 2002US-0414954P.
PR
     09-OCT-2002; 2002US-0417186P.
PR
     09-OCT-2002; 2002US-0417406P.
PR
PR
     23-OCT-2002; 2002US-0420639P.
PR
     28-OCT-2002; 2002US-0421156P.
     31-OCT-2002; 2002US-0422690P.
PR
     01-NOV-2002; 2002US-0423130P.
PR
     05-NOV-2002; 2002US-00423798.
PR
PR
     05-NOV-2002; 2002US-0423798P.
PR
     12-NOV-2002; 2002US-0425453P.
XX
PA
     (CURA-) CURAGEN CORP.
XX
    Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PΙ
    Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PΙ
PΙ
     Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PΙ
     Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
    Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PΙ
PI
     Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
```

```
PΙ
    Rieger DK,
               Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PΙ
    Smithson G,
               Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PΙ
    Zhong H;
XX
    WPI; 2004-081935/08.
DR
DR
    P-PSDB; ADH71646.
XX
PT
    New NOVX polypeptides and nucleic acid molecules useful for preventing or
    treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT
PT
    obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS
    Example 21; SEQ ID NO 541; 1880pp; English.
XX
CC
    The invention relates to a novel isolated polypeptide (NOVX). A
CC
    polypeptide of the invention has cytostatic, immunomodulator,
CC
    neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC
    antilipaemic activity, and may have a use in gene therapy, and as a
    vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC
    any of the 303 fully defined nucleotide sequences given in the
CC
CC
    specification. The polypeptide is useful in the manufacture of a
CC
    medicament for treating a syndrome associated with a human disease. The
    polypeptide, polynucleotide and antibody are useful in diagnosing,
CC
CC
    treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
    Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC
    diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC
    further used as hybridisation probes, in chromosome mapping, tissue
CC
CC
    typing, preventive medicine, and pharmacogenomics. The present sequence
CC .
    encodes a NOVX polypeptide of the invention.
XX
    Sequence 2881 BP; 526 A; 986 C; 868 G; 501 T; 0 U; 0 Other;
SQ
                       97.1%; Score 2619.8; DB 12; Length 2881;
 Query Match
 Best Local Similarity
                      98.9%;
                            Pred. No. 0;
 Matches 2672; Conservative
                             0; Mismatches
                                            22; Indels
                                                                    3;
          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
Qу
            87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146
Db
         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
Qу
            147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206
Db
        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
Qу
            207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266
Db
        181 GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
Qу
            267 GTGCTGCTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326
Db
        241 TGGGTGCGCCAGGTGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
Qу
            327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386
Db
        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
Qу
```

	·Db	387	ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG	446
	Qy	361	GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	420
	Db	447	GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC	506
	Qу	421	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
	Db	507	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	566
	QУ	481	TCCCTGGAGCAGGCATCGTGCCCTGCCGTCCACCGGAGGCCATCCCTCCAGCCGAG	540
	Db	567	TCCCTGGAGCAGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	626
	Qу	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	600
	Db	627	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	686
	Qу	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
	Db	687	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	746
	Qу	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC	720
	Db	747	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTAC	806
	Qy .	721	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCC	780
	Db	807	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	866
,	Qу	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
	Db	867	GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	926
	Qу	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC	897
	Db	927	TGTGAGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC	986
	Qу	898	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	957
	Db	987	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	1046
	Qу	958	CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG	1017
	Db	1047	CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG	1106
	Qу	1018	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1077
	Db	1107	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1166
	Qу	1078	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC	1137
	Db	1167	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC	1226
	Qу	1138	ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1197
	Db	1227	ATCCTCGTTTATTGCCGGAAGAAGAGGGGGCTGGACTCAGATGTGGCTGACTCCATT	1286

	QY .	1198	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1257
	Db	1287	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1346
	Qy	1258	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCAGGGCAGTCTCTGTCCC	1317
	Db	1347	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCC	1403
	Qy	1318	CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1377
٠.	Db	1404	CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1463
	Qу	1378	GGTGGCGCCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1437
	Db	1464	GGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1523
	Qу	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1497
•	Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1583
	Qу	1498	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1557
	Db	1584	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1643
•	Qу	1558	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1617
	Db	1644	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1703
	Qу	1618	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1677
	Db	1704	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1763
	Qy	1678	AGCTGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1737
	Db	1764	AGCTGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1823
	Qy	1738	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1797
•	Db	1824	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1883
	Qу	1798	TGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1854
	Db	1884	TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1943
	Qу	1855	CTGGAGGCCAGTGCCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGA	1914
	Db	1944	CTGGAGGCCAGTGCCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	2003
	Qу	1915	GAGGCCCTCAGCGTGGCTGCCCCAAGCCCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	1974
	Db	2004	GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	2063
	Qy	1975	TGCACCTCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC	2034
	Db	2064		2123

.

.

```
Qу
     2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2094
        Db
     2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183
     Qy
        Db
     2155 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2214
Qу
        2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303
Db
     2215 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274
Qу
        2304 GGCACGCAGCGGTACTTGCACCGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363
Db
     2275 GACCTGGCCTGCAAGCTGTGGCTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334
Qy
        2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423
Db
     2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
Qy
        2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
Db
     2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454
Qу
        Db
     2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543
     2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
Qу
        2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603
Db
     2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCACGCCCCACAGCCATG 2574
Qу
        2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATG 2663
Db
     Qу
        Db
     2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
Qy
        2724 GCAGTGGCTGGCCTGGCCAGCCAGCCTGCCTCTCACAGTGTCGGAGGCTGAGTGC 2783
Db
     2695 TGA 2697
Qу
        \mathbf{H}
     2784 TGA 2786
Db
RESULT 14
TD
   ADH71627 standard; DNA; 2881 BP.
XX
AC
   ADH71627;
XX
DT
   25-MAR-2004
           (first entry)
XX
```

```
DE
     Human gene of the invention NOV21j SEQ ID NO:523.
XX
KW
     ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW
     anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
     vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
KW
     obesity; diabetes; infectious disease; metabolic syndrome X;
KW
     dyslipidaemia.
XX
os
     Homo sapiens.
XX
PN
     WO2003102155-A2.
XX
PD
     11-DEC-2003.
XX
     03-JUN-2003; 2003WO-US017430.
PF
XX
     03-JUN-2002; 2002US-0385120P.
PR
     04-JUN-2002; 2002US-0385784P.
PR
     05-JUN-2002; 2002US-0386041P.
PR
     05-JUN-2002; 2002US-0386047P.
PR
     06-JUN-2002; 2002US-0386376P.
PR
     06-JUN-2002; 2002US-0386453P.
PR
     06-JUN-2002; 2002US-0386864P.
PR
     06-JUN-2002; 2002US-0387016P.
PR
     07-JUN-2002; 2002US-0386796P.
PR
PR
     07-JUN-2002; 2002US-0386816P.
PR
     07-JUN-2002; 2002US-0386931P.
PR
     07-JUN-2002; 2002US-0386942P.
PR
     07-JUN-2002; 2002US-0386971P.
     07-JUN-2002; 2002US-0387262P.
PR
     08-JUN-2002; 2002US-0296960P.
PR
     10-JUN-2002; 2002US-0387400P.
PR
PR
     10-JUN-2002; 2002US-0387535P.
     11-JUN-2002; 2002US-0387610P.
PR
PR
     11-JUN-2002; 2002US-0387625P.
PR
     11-JUN-2002; 2002US-0387634P.
     11-JUN-2002; 2002US-0387668P.
PR
     11-JUN-2002; 2002US-0387696P.
PR
     11-JUN-2002; 2002US-0387702P.
PR
PR
     11-JUN-2002; 2002US-0387836P.
PR
     11-JUN-2002; 2002US-0387859P.
PR
     12-JUN-2002; 2002US-0387933P.
     12-JUN-2002; 2002US-0387934P.
PR
     12-JUN-2002; 2002US-0387960P.
PR
     12-JUN-2002; 2002US-0388022P.
PR
     12-JUN-2002; 2002US-0388096P.
PR
     13-JUN-2002; 2002US-0389123P.
PR
PR
     14-JUN-2002; 2002US-0389118P.
PR
     14-JUN-2002; 2002US-0389120P.
     14-JUN-2002; 2002US-0389144P.
PR
PR
     14-JUN-2002; 2002US-0389146P.
     17-JUN-2002; 2002US-0389729P.
PR
     17-JUN-2002; 2002US-0389742P.
PR
PR
     18-JUN-2002; 2002US-0389884P.
PR
     19-JUN-2002; 2002US-0390006P.
PR
     19-JUN-2002; 2002US-0390209P.
```

PR

21-JUN-2002; 2002US-0390763P.

```
17-JUL-2002; 2002US-0396706P.
PR
     06-AUG-2002; 2002US-0401628P.
PR
     09-AUG-2002; 2002US-0402156P.
PR
     09-AUG-2002; 2002US-0402256P.
PR
     09-AUG-2002; 2002US-0402389P.
PR
     12-AUG-2002; 2002US-0402786P.
PR
     12-AUG-2002; 2002US-0402816P.
PR
     12-AUG-2002; 2002US-0402821P.
PR
PR
     12-AUG-2002; 2002US-0402832P.
PR
     13-AUG-2002; 2002US-0403448P.
     13-AUG-2002; 2002US-0403459P.
PR
PR
     13-AUG-2002; 2002US-0403531P.
     13-AUG-2002; 2002US-0403532P.
PR
     13-AUG-2002; 2002US-0403563P.
PR
     13-AUG-2002; 2002US-0406317P.
PR
     15-AUG-2002; 2002US-0403617P.
PR
     26-AUG-2002; 2002US-0406182P.
PR
     26-AUG-2002; 2002US-0406355P.
PR
PR
     27-AUG-2002; 2002US-0406240P.
     12-SEP-2002; 2002US-0410084P.
PR
     20-SEP-2002; 2002US-0412528P.
PR
     23-SEP-2002; 2002US-0412731P.
PR
     30-SEP-2002; 2002US-0414801P.
PR
     30-SEP-2002; 2002US-0414839P.
PR
PR
     30-SEP-2002; 2002US-0414840P.
PR
     30-SEP-2002; 2002US-0414954P.
     09-OCT-2002; 2002US-0417186P.
PR
     09-OCT-2002; 2002US-0417406P.
PR
     23-OCT-2002; 2002US-0420639P.
PR
     28-OCT-2002; 2002US-0421156P.
PR
     31-OCT-2002; 2002US-0422690P.
PR
     01-NOV-2002; 2002US-0423130P.
PR
PR
     05-NOV-2002; 2002US-00423798.
PR
     05-NOV-2002; 2002US-0423798P.
     12-NOV-2002; 2002US-0425453P.
PR
XX
     (CURA-) CURAGEN CORP.
PA
XX
     Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PΙ
     Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI
                   Gangolli EA, Gerlach VL, Gorman L, Gunther E,
PΙ
     Ettenberg S,
PΙ
     Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
     Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PΙ
                                 Pena CEA, Peyman JA, Raha D,
PΙ
                  Patturajan M,
     Padigaru M,
                                                                 Rastelli L;
                                 Sciore P, Shenoy SG, Shimkets RA;
PI
     Rieger DK,
                 Rothenberg ME,
PΙ
     Smithson G,
                  Spytek KA, Stone DJ, Vernet CAM, Voss EZ,
PΙ
     Zhong H;
XX
DR
     WPI; 2004-081935/08.
     P-PSDB; ADH71628.
DR
XX
PT
     New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT
     treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT
     obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS
     Example 21; SEQ ID NO 523; 1880pp; English.
XX
```

CC The invention relates to a novel isolated polypeptide (NOVX). A polypeptide of the invention has cytostatic, immunomodulator, CC CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and CC antilipaemic activity, and may have a use in gene therapy, and as a vaccine. The polypeptides are encoded by NOVX polynucleotides comprising CC any of the 303 fully defined nucleotide sequences given in the CC specification. The polypeptide is useful in the manufacture of a CC CC medicament for treating a syndrome associated with a human disease. The CC polypeptide, polynucleotide and antibody are useful in diagnosing, CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia, Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are CC CC further used as hybridisation probes, in chromosome mapping, tissue CC typing, preventive medicine, and pharmacogenomics. The present sequence CC encodes a NOVX polypeptide of the invention. XX

SQ Sequence 2881 BP; 525 A; 985 C; 869 G; 502 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2881; Best Local Similarity 98.9%; Pred. No. 0; Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

	Macches	2012	z, conservative	Ο,	PILSMacches	22,	Inders	,	Gaps	٠,
Q	У	1	ATGGCCGTCCGGCCCGGC							60
D	b	87	ATGGCCGTCCGGCCCGGCC							146
Q	У	61	CGCGGCTCGGGTGCCCAG							120
D	b	147	CGCGGCTCGGGTGCCCAG							206
Q	У	121	GACCTGCTTCCCCACTTC							180
D	b ,	207	GACCTGCTTCCCCACTTC							266
Q	у .	181	GTGCTGCTTGTGTGCAAG							240
D	b	267	GTGCTGCTTGTGTGCAAG							326
Q	У	241	TGGGTGCGCCAGGTGGAC							300
D	b	327	TGGGTGCGCCAGGTGGAC						 GAGCCG	386
Q	У	301	ACCATGGAGGTCCGCATT							360
D	b	387	ACCATGGAGGTCCGCATT							446
Q	<i>,</i> У	361	GAATACTGGTGCCAGTGC							420
D	b	447	GAATACTGGTGCCAGTGC							506
Q	У	421	TACATCCGCATAGCCAGA'	TTGC	GCAAGAACTTC	GAGCAGG	AGCCGCTGG	CCAAG	GAGGTG	480
D	b	507	TACATCCGCATAGCCAGA							566
Q	v	481	TCCCTGGAGCAGGGCATC	GTGC	rgccctgccgt	CCACCGG	AGGGCATCC	CTCCA	GCCGAG	540
D	_		TCCCTGGAGCAGGGCATC		11111111111	111111	ППППП			
_	~	307	1000100/100/100000110	0100.	10000100001	20110000		-1	2200.10	525

Qу	. 541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCCGTCCCTGGACCCCAATGTATACATC	600
Db	627		686
Qy	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
Db	687	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	746
Qу	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC	720
Db	747	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTCATCGTCTAC	806
Qу	721	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCC	780
Db	807	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCC	866
Qy	. 781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
Db	867	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	926
Qy	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC	897
Db	927	TGTGAGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGTTGTCTCTGTGGACGGC	986
Qу	898	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	957
Db	987	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	1046
Qу	958	CGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGGGAGG	1017
Db	1047	CGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGGGAGG	1106
Qу	1018	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1077
Db	1107	GACACCCGCAACTGTACCAGTGACCTCTGTGTGCACAGTGCTTCTGGCCCTGAGGACGTG	1166
Qу	1078	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC	1137
Db	1167	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC	1226
Qу		ATCCTCGTTTATTGCCGGAAGAAGGAGGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1197
Db		ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1286
Qу	1198	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1257
Db	1287	CTCACCTCAGGCTTCCAGCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1346
Qу	1258	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCC	1317
Db	1347		1403
Qу	1318	CGGCAGGATGGGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1377
Db	1404		1463

Qу	1378	GGTGGCGCCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1437
Db	1464	GGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1523
Qy .	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1497
Db	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1583
Qу	1498	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1557
Db	1584	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1643
Qу	1558	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1617
Db	1644		1703
QУ	1618	AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT	1677
Db	1704		1763
Qу	1678	AGCTGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1737
Db	1764	AGCTGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT	1823
Qу	1738	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1797
Db	1824	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1883
Qу	1798	TGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1854
Db	1884	TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1943
Qу	1855	CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	1914
Db	1944	CTGGAGGCCAGTGCCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	2003
Qу	1915	GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	1974
Db	2004	GAGGCCCTCAGCGTGGCTGCCCCCAAGCCTCCTCTTTTTTTT	2063
Qу	1975	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2034
Db	2064	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2123
Qу	2035	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC	2094
Db	2124	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC	2183
Qу	2095	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	2154
Db	2184	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCACCACGATGTGCCCAGCTCC	2243
Qу	2155	CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT	2214
Db	2244	CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT	2303
Qу	2215	GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT	2274

```
2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363
Db
      2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334
Qу
          2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423
Db
      2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
Qy
          2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483
Db
      2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454
Qу
          2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543
Db
      2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
Qу
          2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603
Db
      2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2574
Qy
          2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATG 2663
Db
      Qу
          Db
      2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
Qу
          2724 GCAGTGGCTGGACTGGCCAGCCAGACGCTGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
Db
      2695 TGA 2697
Qу
          | | |
      2784 TGA 2786
Db
RESULT 15
ADH71639
   ADH71639 standard; DNA; 2881 BP.
XX
AC
   ADH71639;
XX
   25-MAR-2004 (first entry)
DT
XX
   Human gene of the invention NOV21p SEQ ID NO:535.
DE
XX
   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW
   anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW
   vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
   obesity; diabetes; infectious disease; metabolic syndrome X;
KW
   dyslipidaemia.
KW
XX
os
   Homo sapiens.
XX
PN
   WO2003102155-A2.
XX
PD
   11-DEC-2003.
```

```
XX
     03-JUN-2003; 2003WO-US017430.
PF
XX
     03-JUN-2002; 2002US-0385120P.
PR
     04-JUN-2002; 2002US-0385784P.
PR
     05-JUN-2002; 2002US-0386041P.
PR
     05-JUN-2002; 2002US-0386047P.
PR
     06-JUN-2002; 2002US-0386376P.
PR
     06-JUN-2002; 2002US-0386453P.
PR
     06-JUN-2002; 2002US-0386864P.
PR
     06-JUN-2002; 2002US-0387016P.
PR
PR
     07-JUN-2002; 2002US-0386796P.
     07-JUN-2002; 2002US-0386816P.
PR
     07-JUN-2002; 2002US-0386931P.
PR
     07-JUN-2002; 2002US-0386942P.
PR
     07-JUN-2002; 2002US-0386971P.
PR
PR
     07-JUN-2002; 2002US-0387262P.
     08-JUN-2002; 2002US-0296960P.
PR
PR
     10-JUN-2002; 2002US-0387400P.
     10-JUN-2002; 2002US-0387535P.
PR
     11-JUN-2002; 2002US-0387610P.
PR
     11-JUN-2002; 2002US-0387625P.
PR
PR
     11-JUN-2002; 2002US-0387634P.
PR
     11-JUN-2002; 2002US-0387668P.
PR
     11-JUN-2002; 2002US-0387696P.
     11-JUN-2002; 2002US-0387702P.
PR
     11-JUN-2002; 2002US-0387836P.
PR
PR
     11-JUN-2002; 2002US-0387859P.
     12-JUN-2002; 2002US-0387933P.
PR
     12-JUN-2002; 2002US-0387934P.
PR
     12-JUN-2002; 2002US-0387960P.
PR
PR
     12-JUN-2002; 2002US-0388022P.
₽R
     12-JUN-2002; 2002US-0388096P.
PR
     13-JUN-2002; 2002US-0389123P.
     14-JUN-2002; 2002US-0389118P.
PR
PR
     14-JUN-2002; 2002US-0389120P.
     14-JUN-2002; 2002US-0389144P.
PR
     14-JUN-2002; 2002US-0389146P.
PR
     17-JUN-2002; 2002US-0389729P.
PR
PR
     17-JUN-2002; 2002US-0389742P.
     18-JUN-2002; 2002US-0389884P.
PR
PR
     19-JUN-2002; 2002US-0390006P.
     19-JUN-2002; 2002US-0390209P.
PR
     21-JUN-2002; 2002US-0390763P.
PR
     17-JUL-2002; 2002US-0396706P.
PR
     06-AUG-2002; 2002US-0401628P.
PR
     09-AUG-2002; 2002US-0402156P.
PR
PR
     09-AUG-2002; 2002US-0402256P.
PR
     09-AUG-2002; 2002US-0402389P.
     12-AUG-2002; 2002US-0402786P.
PR
     12-AUG-2002; 2002US-0402816P.
PR
     12-AUG-2002; 2002US-0402821P.
PR
PR
     12-AUG-2002; 2002US-0402832P.
     13-AUG-2002; 2002US-0403448P.
PR
PR
     13-AUG-2002; 2002US-0403459P.
PR
     13-AUG-2002; 2002US-0403531P.
     13-AUG-2002; 2002US-0403532P.
PR
```

```
13-AUG-2002; 2002US-0403563P.
PR
     13-AUG-2002; 2002US-0406317P.
PR
     15-AUG-2002; 2002US-0403617P.
PR
     26-AUG-2002; 2002US-0406182P.
PR
PR
     26-AUG-2002; 2002US-0406355P.
     27-AUG-2002; 2002US-0406240P.
PR
     12-SEP-2002; 2002US-0410084P.
PR
     20-SEP-2002; 2002US-0412528P.
PR
     23-SEP-2002; 2002US-0412731P.
PR
     30-SEP-2002; 2002US-0414801P.
PR
PR
     30-SEP-2002; 2002US-0414839P.
PR
     30-SEP-2002; 2002US-0414840P.
     30-SEP-2002; 2002US-0414954P.
PR
     09-OCT-2002; 2002US-0417186P.
PR
     09-OCT-2002; 2002US-0417406P.
PR
     23-OCT-2002; 2002US-0420639P.
PR
     28-OCT-2002; 2002US-0421156P.
PR
PR
     31-OCT-2002; 2002US-0422690P.
PR
     01-NOV-2002; 2002US-0423130P.
     05-NOV-2002; 2002US-00423798.
PR
     05-NOV-2002; 2002US-0423798P.
PR
     12-NOV-2002; 2002US-0425453P.
PR
XX
PA
     (CURA-) CURAGEN CORP.
XX
PΙ
     Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
     Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PΙ
ΡI
     Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
     Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PΙ
    Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PΙ
     Padigaru M,
                                Pena CEA, Peyman JA, Raha D, Rastelli L;
PΙ
                  Patturajan M,
PΙ
                                Sciore P, Shenoy SG, Shimkets RA;
     Rieger DK, Rothenberg ME,
PΙ
                 Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
     Smithson G,
PI
     Zhong H;
XX
DR
    WPI; 2004-081935/08.
DR
     P-PSDB; ADH71640.
XX
    New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT
PT
     treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT
     obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS
     Example 21; SEQ ID NO 535; 1880pp; English.
XX
CC
     The invention relates to a novel isolated polypeptide (NOVX). A
CC
     polypeptide of the invention has cytostatic, immunomodulator,
CC
     neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC
     antilipaemic activity, and may have a use in gene therapy, and as a
CC
     vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
     any of the 303 fully defined nucleotide sequences given in the
CC
CC
     specification. The polypeptide is useful in the manufacture of a
CC
     medicament for treating a syndrome associated with a human disease. The
CC
     polypeptide, polynucleotide and antibody are useful in diagnosing,
CC
     treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC
     Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC
     diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC
     further used as hybridisation probes, in chromosome mapping, tissue
```

CC typing, preventive medicine, and pharmacogenomics. The present sequence CC encodes a NOVX polypeptide of the invention. XX Sequence 2881 BP; 525 A; 985 C; 869 G; 502 T; 0 U; 0 Other; SO 97.1%; Score 2619.8; DB 12; Length 2881; Best Local Similarity 98.9%; Pred. No. 0; 0; Mismatches 9; Gaps 3; Matches 2672; Conservative 22; Indels 1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60 Qу 87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146 Db 61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120 Qу 147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206 Db 121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180 Qу 207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266 Db 181 GTGCTGCTGTGTGCAAGGCCGTGCCCGCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240 Qу 267 GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326 Db 241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300 Qу 327 TGGGTGCGCCAGGTGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386 Db 301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360 Qу 387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446 Db 361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420 Qу 447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506 Db 421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480 Qy 507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566 Db 481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540 Qy 567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626 Db 541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600 Qy 627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686 Db 601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660 Qy 687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746 Db

661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTTCATCGTCTAC 720

Qy

Db

Qу	721	$\tt GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCCAGCTGTGGGGCGCCAGCTGTGGGGCGCCAGCTGTGGGGCGCCAGCTGTGGGGCGCCAGCTGTGGGGCGCCAGCTGTGGGGCGCCAGCTGTGGGGCGCCAGCTGTGGGGCGCCAGCTGTGGGGCGCCAGCTGTGGGGCGCCAGCTGTGGGGCGCCAGCTGTGGGGCGCCAGCTGTGGGGCGCCAGCTGTGGGGCGCCAGCTGTGGGGCGCCAGCTGTGGGGCGCCAGCTGTGGGGCGCCAGCTGTGGGGCGCCAGCTGTGGGGCGCCAGCTGGGGGCGCCAGCTGGGGGCGCCAGCTGGGGGCGCCAGCTGGGGCGCAGCGCGCAGCGCGCAGCGCGCAGCGCGCAGCGCGCAGCGCGCAGCGCGCAGCGCGCAGCGCAGGGGCGCAGCGCGCAGCGGCG$	780
Db·	807		866
Qу	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
Db	867		926
Qу	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC	897
Db	927	TGTGAGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGTGTCTCTGTGGACGGC	986
Qу	898	AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGAGC	957
Db	987		1046
Qу	958	CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG	1017
Db	1047	CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG	1106
Qу	1018	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1077
Db	1107	GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG	1166
Qу	1078	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC	1137
Db	1167	GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC	1226
Qу	1138	ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1197
Db	1227	ATCCTCGTTTATTGCCGGAAGAAGGAGGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT	1286
Qy	1198	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1257
Db	1287	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG	1346
Qу	1258	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCAGGGCAGTCTCTGTCCC	1317
Db	1347	CTCACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCC	1403
Qу	1318	CGGCAGGATGGGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1377
Db	1404	CGGCAGGATGGGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG	1463
Qу	1378	GGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1437
Db	1464	GGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC	1523
Qy	1438	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1497
Db ·	1524	TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC	1583
Qу	1498	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1557
Db	1584	TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC	1643
Qу	1558	$\tt CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC$	1617

.

Db	1644	CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC	1703
Qy			1677
Db ·		${\tt AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT} \\ {\tt AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT} \\ {\tt CAGCTGTGGACCCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT} \\ {\tt CAGCTGTGGACCCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT} \\ {\tt CAGCTGTGGACCCCCTGGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT} \\ {\tt CAGCTGTGGACCCCCTGGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT} \\ {\tt CAGCTGTGGACCCCCTGGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT} \\ {\tt CAGCTGTGGACCCCCTGGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT} \\ {\tt CAGCTGTGGACCCCTGGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT} \\ {\tt CAGCTGTGGACCACTGT} \\ {\tt CAGCTGTGGACCCCTGGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT} \\ {\tt CAGCTGTGGACCACTGT} \\ {\tt CAGCTGTGACCCGGCCAGTCATCCTGGCTATGGACCACTGT} \\ {\tt CAGCTGTGGACCACTGT} \\ {\tt CAGCTGTGGACTGTGGACCACTGT} \\ {\tt CAGCTGTGGACTGTGGACCACTGT} \\ {\tt CAGCTGTGGACTGTGACTGTGGACTGTGGACTGTGACTGTGGACTGTGACTGTGGACTGTGACTGAC$	1763 1737
Db	1764		1823
Qу	1738	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1797
Db	1824	GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC	1883
QУ	1798	TGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1854
Db	1884	TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG	1943
Qy	1855	CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	1914
Db	1944	CTGGAGGCCAGTGCCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA	2003
Qу	1915	GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC	1974
Dр	2004	GAGGCCCTCAGCGTGGCTGCCCCAAGCGCCTCAGGCTGCTTCTGTTTGCGCCGGTGGCC	2063
Qу	1975	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2034
Db	2064	TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC	2123
Qу	2035	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC	2094
Db	2124	AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC	2183
Qу	2095	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	2154
Db	2184	CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC	2243
Qу	2155	CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT	2214
Db	2244	CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT	2303
Qу	2215	GGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT	2274
Db	2304	GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT	2363
Qу	2275	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC	2334
Db	2364	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC	2423
Qy .	2335	AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2394
Db	2424	AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG	2483
Qy .	2395	GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG	2454

Db	2484	GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543
Qу	2455	ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
Db	2544	ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603
Qy	2515	AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATG 2574
Db	2604	AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATG 2663
Qy	2575	ATCCTCAACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC
Db	2664	ATCCTCAACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC
Qy	2635	GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
Db	2724	GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
Qу	2695	TGA 2697
Db	2784	TGA 2786

Search completed: March 5, 2005, 22:58:58
Job time: 1381 secs

GenCore version 5.1.6 Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2005, 19:26:23; Search time 424.67 Seconds

(without alignments)

10391.692 Million cell updates/sec

Title: US-10-624-932-1 COPY 46 2742

Perfect score: 2697

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

Q.

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*

2: /cgn2 6/ptodata/1/ina/5B COMB.seq:*

3: /cgn2 6/ptodata/1/ina/6A COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2 6/ptodata/1/ina/PCTUS COMB.seq:*

6: /cgn2 6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		*				
Result		Query				
No.	Score	Match	Length	DB	ID	Description
1	2252.2	83.5	3014	2	US-08-808-982-1	Sequence 1, Appli
2	2252.2	83.5	3014	3	US-09-306-902A-1	Sequence 1, Appli
3	1552.4	57.6	1787	2	US-08-808-982-2	Sequence 2, Appli
4	1552.4	57.6	1787	3	US-09-306-902A-2	Sequence 2, Appli
5	841.4	31.2	2831	2	US-08-808-982-3	Sequence 3, Appli
6	841.4	31.2	2831	3	US-09-306-902A-3	Sequence 3, Appli
7	833.6	30.9	3008	4	US-09-949-016-4794	Sequence 4794, Ap
8	487	18.1	2736	4	US-09-969-532-9	Sequence 9, Appli
9	487	18.1	3411	4	US-09-969-532-33	Sequence 33, Appl
10	467	17.3	2703	4	US-09-969-532 - 11	Sequence 11, Appl
11	460.2	17.1	2694	4	US-09-969-532-13	Sequence 13, Appl

```
439.8
  12
               16.3
                     2661 4
                             US-09-969-532-15
                                                       Sequence 15, Appl
                     349 4
  13 . 323
               12.0
                                                       Sequence 345, App
                             US-09-471-276-345
              10.9 1968 4
  14
       293.2
                             US-09-969-532-31
                                                       Sequence 31, Appl
               10.9
  15
       293.2
                     2001 4
                             US-09-969-532-29
                                                       Sequence 29, Appl
  16
       293.2
               10.9 2010 4 US-09-969-532-27
                                                       Sequence 27, Appl
  17
       293.2
              10.9 2043 4 US-09-969-532-25
                                                       Sequence 25, Appl
  18
       284.8
               10.6 1659 4
                             US-09-969-532-7
                                                       Sequence 7, Appli
                             US-09-969-532-5
  19
       284.8
               10.6 1692 4
                                                       Sequence 5, Appli
              10.6 1701 4
  20
       284.8
                             US-09-969-532-3
                                                       Sequence 3, Appli
                     1734 4
  21
       284.8
               10.6
                             US-09-969-532-1
                                                       Sequence 1, Appli
  22
С
         269
               10.0
                    771 1
                             US-08-253-155A-17
                                                       Sequence 17, Appl
c 23
             6.7 19326 4
       181.4
                             US-09-949-016-16776
                                                       Sequence 16776, A
  24
       123.4
               4.6 114139 4
                             US-09-949-016-16536
                                                       Sequence 16536, A
  25
        119
               4.4
                      305 2 US-08-808-982-4
                                                       Sequence 4, Appli
  26
        119
               4.4
                      305 3 US-09-306-902A-4
                                                       Sequence 4, Appli
                      601 4
  27
        80.4
               3.0
                             US-09-949-016-169999
                                                       Sequence 169999,
                      966 4
  28
        80.2
               3.0
                             US-09-969-532-23
                                                       Sequence 23, Appl
  29
        75.8
               2.8 1008 4 US-09-969-532-19
                                                       Sequence 19, Appl
c 30
        73.8
               2.7 601 4 US-09-949-016-170033
                                                       Sequence 170033,
c 31
        66.4
               2.5 601 4 US-09-949-016-170040
                                                       Sequence 170040,
c 32
        66.4
               2.5 601 4 US-09-949-016-170041
                                                       Sequence 170041,
  33
        63.6
               2.4 5784 4 US-09-949-016-462
                                                       Sequence 462, App
c 34
        63
               2.3
                   601 4 US-09-949-016-170000
                                                       Sequence 170000,
                     601 4
  35
          63
               2.3
                                                       Sequence 170034,
                             US-09-949-016-170034
  36
        60.2
               2.2
                     2820 4
                             US-09-854-845-15
                                                       Sequence 15, Appl
  37
        60.2
               2.2 2865 4
                             US-09-854-845-13
                                                       Sequence 13, Appl
  38
        60.2
               2.2 3105 4 US-09-854-845-5
                                                       Sequence 5, Appli
  39
        60.2
               2.2 3150 4 US-09-854-845-1
                                                       Sequence 1, Appli
  40
        60.2
               2.2 3237 4 US-09-854-845-7
                                                       Sequence 7, Appli
  41
        60.2
               2.2 3282 4
                             US-09-854-845-3
                                                       Sequence 3, Appli
  42
        60.2
               2.2 3411 4
                             US-09-854-845-11
                                                       Sequence 11, Appl
                   3456 4
  43
        60.2
               2.2
                             US-09-854-845-9
                                                       Sequence 9, Appli
                   4074 4
  44
        60.2
               2.2
                             US-09-854-845-17
                                                       Sequence 17, Appl
  45
        56.4
               2.1
                     5178 4 US-09-949-016-5241
                                                       Sequence 5241, Ap
```

ALIGNMENTS

```
RESULT 1
US-08-808-982-1
; Sequence 1, Application US/08808982
; Patent No. 5939271
  GENERAL INFORMATION:
;
    APPLICANT: Tessier-Lavigne, Marc
;
    APPLICANT: Leonardo, E. David
;
    APPLICANT: Hink, Lindsay
;
    APPLICANT: Masu, Masayuki
;
    APPLICANT: Kazuko, Keino-Masu
;
    TITLE OF INVENTION: Netrin Receptors
    NUMBER OF SEQUENCES: 8
;
;
    CORRESPONDENCE ADDRESS:
;
      ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
      STREET: 268 BUSH STREET, SUITE 3200
;
      CITY: SAN FRANCISCO
;
      STATE: CALIFORNIA
;
      COUNTRY: USA
```

```
ZIP: 94104
   COMPUTER READABLE FORM:
     MEDIUM TYPE: Floppy disk
     COMPUTER: IBM PC compatible
     OPERATING SYSTEM: PC-DOS/MS-DOS
     SOFTWARE: PatentIn Release #1.0, Version #1.30
   CURRENT APPLICATION DATA:
     APPLICATION NUMBER: US/08/808,982
     FILING DATE:
     CLASSIFICATION: 530
   ATTORNEY/AGENT INFORMATION:
     NAME: OSMAN, RICHARD A
     REGISTRATION NUMBER: 36,627
;
     REFERENCE/DOCKET NUMBER: UC96-217
   TELECOMMUNICATION INFORMATION:
     TELEPHONE: (415) 343-4341
     TELEFAX: (415) 343-4342
  INFORMATION FOR SEQ ID NO: 1:
   SEQUENCE CHARACTERISTICS:
     LENGTH: 3014 base pairs
     TYPE: nucleic acid
     STRANDEDNESS: double
     TOPOLOGY: linear
   MOLECULE TYPE: cDNA
US-08-808-982-1
 Query Match
                    83.5%; Score 2252.2; DB 2; Length 3014;
 Best Local Similarity 89.7%; Pred. No. 0;
 Matches 2419; Conservative
                         0; Mismatches 278; Indels
                                                  0; Gaps
                                                           0;
         1 ATGGCCGTCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
Qу
          1 ATGGCCGTCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60
Db
        61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
Qу
          61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120
Db
       121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
Qу
          121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180
Db
       181 GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
Qу
          181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240
Db
       241 TGGGTGCGCCAGGTGACCACGTGATCGAGCGCAGCAGACGGGAGCAGTGGGCTGCCC 300
Qу
          Db
       301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
Qу
           Db
       301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360
       361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
Qy
          361 GAATACTGGTGCCAGTGTGGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420
Db
```

Qу	421	TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG	480
Db	421	${\tt TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG}$	480
Qу	481	TCCCTGGAGCAGGGCATCGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG	540
Db	481	${\tt TCACTGGAGCAAGGCATTGTACCTTGTCGCCCCCAGAAGGAATCCCCCCAGCTGAG}$	540
Qу	541	GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC	600
Db	541	GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCGATCCCAATGTGTACATC	600
Qу	601	ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC	660
Db	601	ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC	660
Qу	661	TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC	720
Db	661	TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT	720
Qу	721	GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC	780
Db	721	GTGAACGGTGGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCGT	780
Qу	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC	840
Db	781	GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC	840
Qу	841	TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC	900
Db	841	TGTGAGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGGAGC	900
Qу	901	TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT	960
Db	901	TGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGCCGC	960
Qу		GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC	
Db	961	GAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGGTGCTGACCTGGAC	1020
Qу	1021	ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC	1080
Db	1021	ACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTGGCT	1080
Qу	1081	CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTCATC	1140
Db	1081	CTCTACATCGGCCTTGTCGCTGTGGCTGTGCCTCTTCTTGCTGTTGCTGGCCCTTGGA	1140
Qу	1141	CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC	1200
Db	1141	CTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATCCTC	1200
Qу	1201	ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC	1260
Db	1201	ACCTCGGGCTTCCAGCCTGTCAGCCCCAGCAAAGCAGACAACCCCCACCTGCTC	1260